Treatment effects of fixed functional appliances in patients with Class II malocclusion: a systematic review and meta-analysis

Zymperdikas, Vasileios F ; Koretsi, Vasiliki ; Papageorgiou, Spyridon N ; Papadopoulos, Moschos A

Abstract: OBJECTIVE To assess the treatment effects of fixed functional appliances (FFAs) in treated versus untreated Class II patients by means of lateral cephalometric radiographs. SEARCH METHODS Unrestricted electronic search of 18 databases and additional manual searches up to October 2014. SELECTION CRITERIA Prospective randomized and non-randomized controlled trials reporting on cephalometric angular measurements of Class II patients treated with FFAs and their matched untreated controls. DATA COLLECTION AND ANALYSIS Skeletal, dental, and soft tissue cephalometric data were annualized and stratified according to the time of evaluation in effects. Following risk of bias evaluation, the mean differences (MDs) and 95 % confidence intervals (CIs) were calculated with random-effects models. Patient- and appliance-related subgroup analyses and sensitivity analyses were performed with mixed-effects models. RESULTS Nine studies were included (244 patients; mean age: 13.5 years and 174 untreated controls; mean age: 12.8 years) reporting on cephalometric effects directly after the removal of FFAs. FFAs were found to induce a small reduction of SNA angle (MD = -0.83 degree/year, 95 % CI: -1.17 to -0.48), a small increase of SNB angle (MD = 0.87 degree/year, 95 % CI: 0.30-1.43), and moderate decrease of ANB angle (MD = -1.74 degree/year, 95 % CI: -2.50 to -0.98) compared to untreated Class II patients. FFA treatment resulted in significant dentoalveolar and soft tissue changes. Several patient- or appliance-related factors seem to affect the treatment outcome. Long-term effectiveness of FFAs could not be assessed due to limited evidence. CONCLUSIONS According to existing evidence, FFAs seem to be effective in improving Class II malocclusion in the short term, although their effects seem to be mainly dentoalveolar rather than skeletal.

DOI: https://doi.org/10.1093/ejo/cjv034

Treatment effects of fixed functional appliances in patients with Class II malocclusion: A systematic review and meta-analysis

Vasileios F. Zymperdikas*, Vasiliki Koretsi**, Spyridon N. Papageorgiou***,****,***** and Moschos A. Papadopoulos******

*Dental Unit, Medical Company, 71st Airborne Brigade, Nea Santa, Greece, **Department of Orthodontics, School of Dentistry, University Medical Centre Regensburg, Regensburg, Germany, ***Department of Orthodontics, ****Department of Oral Technology, and *****Clinical Research Unit 208, University of Bonn, Germany, ******Department of Orthodontics, School of Health Sciences, Faculty of Dentistry, Aristotle University of Thessaloniki, Thessaloniki, Greece.

Running title - Fixed functional appliances in Class II treatment

Conflict of interest - The authors declare no funding or conflicts of interest.

Acknowledgements

We thank the following author for providing clarifications for the systematic review: O. Dalci (Discipline of Orthodontics, Faculty of Dentistry, University of Sydney, Sydney, Australia).

***Correspondence to*:** Prof. Moschos A. Papadopoulos, Chairman and Program Director, Department of Orthodontics, School of Health Sciences, Faculty of Dentistry, Aristotle University of Thessaloniki, GR-54124 Thessaloniki, Greece; Tel.: +30-2310-999482, +30-2310-999556; Fax: +30-2310-999549; e-mail: mikepap@dent.auth.gr

Treatment effects of fixed functional appliances in patients with Class II malocclusion: A systematic review and meta-analysis

SUMMARY

OBJECTIVE: To assess the treatment effects of fixed functional appliances (FFAs) in treated versus untreated Class-II patients by means of lateral cephalometric radiographs.

SEARCH METHODS: Unrestricted electronic search of 18 databases and additional manual searches up to October 2014.

SELECTION CRITERIA: Prospective randomized and non-randomized controlled trials reporting on cephalometric angular measurements of Class-II patients treated with FFAs and their matched control patients.

DATA COLLECTION AND ANALYSIS: Skeletal, dental, and soft tissue cephalometric data were annualized and stratified according to the time of evaluation in effects. Following Risk of Bias evaluation, the Mean Differences (MDs) and 95% Confidence Intervals (CIs) were calculated with random-effects models. Patient- and appliance-related subgroup analyses, and sensitivity analyses were performed with mixed-effects models.

RESULTS: Nine studies were included (418 patients; mean age: 13.2 years) reporting on cephalometric effects directly after the removal of FFAs. FFAs were found to induce a small reduction of SNA angle (MD=-0.83º/year, 95% CI: -1.17 to -0.48), a small increase of SNB angle (MD=0.87º/year, 95% CI: 0.30 to 1.43), and moderate decrease of ANB angle (MD = -1.74º/year, 95% CI: -2.50 to -0.98) compared to untreated Class-II patients. FFA treatment resulted in significant dentoalveolar and soft-tissue changes. Several patient- or appliance-related factors seem to affect the treatment outcome. Long-term effectiveness of FFAs could not be assessed due to limited evidence.

CONCLUSIONS: According to existing evidence, FFAs seem to be effective in improving Class-II malocclusion in the short-term, although their effects seem to be mainly dentoalveolar rather than skeletal.

The authors declare no conflicts of interest.

Running title: Fixed functional appliances in Class II treatment

Introduction

Rationale

Class II malocclusion is a frequent condition and a rather common reason for patients seeking orthodontic treatment (1, 2). The development of Class II malocclusion could be attributed to several factors, hence their accurate diagnosis is important for the selection of the corresponding treatment plan. Among these factors, mandibular retrognathism is considered a dominant one (1, 3). In this respect, mandibular advancement through the use of functional appliances is considered as a popular treatment approach in growing individuals.

While some researchers posit favourable treatment outcomes based on mandibular growth, attributed either as a mandibular length augmentation or effective condyle growth (4–7), others dispute the magnitude of these effects (8, 9). Furthermore, a restriction effect on the maxilla has been likewise supported by some researchers (10–14) and questioned by others (5, 15). Moreover, existing evidence indicates that the dentoalveolar changes produced by functional treatment outweigh the skeletal changes attained (13, 15, 16, 17).

Functional appliances can be categorised into either removable or fixed ones (FFAs). An important discriminating factor between them is the need for patient compliance, which is considered to as a possible influence on the treatment outcomes (18, 19). As a consequence, it is essential to assess those two types of functional appliances separately, in order to investigate their clinical effectiveness and to reveal any existing differences. Previously published systematic reviews on the subject presented methodological limitations (9, 20–30). The current systematic review on fixed functional appliances supplements a previously published systematic review focused exclusively on removable functional appliances (17).

Objectives

This study aims to summarize current evidence only from randomized controlled trials (RCTs) and prospective controlled clinical trials (pCCTs) assessing by means of lateral cephalometric radiographs the clinical effectiveness of FFAs for the treatment of patients with Class II malocclusion in comparison with untreated individuals, as well as to identify any factors affecting the treatment outcomes.

Materials and Methods

Protocol and registration

The protocol for the present systematic review was constructed *a priori* according to the Cochrane Handbook for Systematic Reviews of Interventions 5.1.0 (31) and is available upon request (no registration was performed). The systematic review is reported on the basis of the PRISMA statement (32) and its extension for abstracts (33).

Information sources and search

Eighteen electronic databases were systematically and unrestrictedly searched up to November 2011 and updated in October 2014. In an effort to screen as many eligible articles as possible and not to miss any pertinent studies, the search strategy covered initially all functional appliances used for Class II treatment (both removable and fixed), and then was cautiously limited to FFAs. MESH terms and the respective keywords were used properly to fit each database (Supplementary table 1). The search strategy included no limitations concerning language, publication year or status. The reference lists of the included trials and relevant reviews were manually searched as well. Grey literature was also assessed through proper registers and databases. When considered necessary, authors were contacted for complementary data or clarifications. The search was performed independently by two authors (VFZ and VK).

Eligibility criteria and study selection

The eligibility criteria were pre-determined (Table 1). A study was considered eligible when it reported on at least one treatment arm with a FFA and simultaneously all of the inclusion and none of the exclusion criteria were fulfilled. In order to investigate only the effects of FFAs, data concerning any previous or subsequent phases with fixed appliances was not included, since fixed appliances are likely to alter the effects caused by functional treatment (30). After the elimination of duplicates, the decision for the selection was made by taking into consideration the title, abstract and, when it was considered necessary, the full-text of the respective articles. Multiple reports pertinent to the same trial/patient cohort were grouped together. When the same trial was published in various languages, the English version was preferred. Finally, articles including at least one treatment arm with FFAs were selected.

Data collection process and data items

Data was extracted independently on pre-defined and piloted forms by two authors (VFZ and VK). Any ambiguities were resolved after discussion with the last author (MAP). In an effort to investigate the clinical effects of FFAs in Class II treatment, only angular cephalometric measurements on lateral cephalometric analyses were considered as primary outcomes, due to the fact that linear measurements are prone to magnification bias (34, 35). Due to the variability of the terms used among the authors for identical variables, all equivalent terms pertaining to the same variable were grouped (Supplementary table 2) and one term was used throughout the review.

If the same variable was reported in at least two included trials, the respective data were extracted and categorised as skeletal (sagittal and vertical), dental, and soft tissue variables. Reported outcomes were stratified based on the time of evaluation in effects: (a) after the removal of the corresponding FFA and (b) after the retention phase.

Several factors were determined *a priori* to be investigated via subgroup analyses for their possible effect on the FFA treatment outcomes, if at least five studies reported on the corresponding factor. These factors were divided to patient-related (i.e. specific patients' characteristics) and appliance-related (i.e. specific features of the appliance design or the treatment plan). The patient-related factors involved (a) patient gender ratio (male patients/female patients), (b) patients' skeletal growth stage (pre-peak and peak or post-peak), and (c) patients' growth pattern (horizontal, vertical, or average). The discrimination between the various stages of skeletal maturation relied exclusively on the cervical vertebral maturation index or on hand-wrist radiographs, which are considered as efficient methods for the identification of skeletal age (36, 37). Studies providing data concerning only the chronological and/or dental age of the patients were excluded from the evaluation of skeletal maturity, since these methods are not considered reliable for the discrimination of skeletal growth stages (38, 39). If the growth pattern was clearly stated in the included articles, then it was categorised accordingly. When no such data was available, the discrimination was based on the mean values of either the SN-ML or FH-ML angles, as reported on the patients' baseline characteristics.

Further, the appliance-related factors included (a) appliance used (i.e. the exact type and design of the respective appliance) and (b) construction bite (single step vs. stepwise mandibular advancement).

Risk of bias in individual studies

The Cochrane Collaboration's tool (31) and a modified Downs and Black checklist (40) were used to assess the risk of bias in RCTs and pCCTs, respectively. The overall risk of bias was judged as "high", "low" or "unclear" for randomized studies, while serious methodological limitations were judged to exist when a pCCT collected less than 17 points on the modified checklist (41).

Risk of bias across studies

If a sufficient number of trials were identified (n >10) reporting biases (small-study effects or publication bias) were planned to be assessed through the inspection of a contour-enhanced funnel plot (42), Begg's rank correlation test (43) and Egger's weighted regression test (44). If the tests hinted towards the existence of publication bias, the Duval & and Tweedie's trim and fill procedure (45) was planned to be performed.

The overall quality of evidence (confidence in effect estimates) for each of the main outcomes was rated by using the Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) approach (46).

The minimal clinical important, large, and very large effects were conventionally defined (47) as half, one, and two standard deviations of the Caucasian cephalometric norm plus 1°, respectively, to allow for method error. The optimal information size (i.e. required meta-analysis sample size) was calculated for each outcome independently for $\alpha = 5\%$ and $\beta = 20\%$.

The search strategy, study selection, data extraction, and within-studies risk of bias assessment were performed independently without blinding (48) by two review authors (VFZ and VK), across-studies risk of bias assessment by a third author (SNP), and any disagreement was resolved by discussion with the last author (MAP). Authors were contacted when necessary to resolve ambiguities or provide complementary data. Inter-reviewer agreement for the three stages was evaluated with the un-weighted Cohen's kappa (49).

Summary measures and synthesis of results

Data were summarised and considered suitable for pooling if similar control groups of untreated Class II patients were used and if the same cephalometric angular outcomes were reported. In an effort to account for the different follow-up periods of the included studies, the treatment and the observational changes were annualized. Mean Differences (MDs) and their corresponding 95% Confidence Intervals (CIs) were calculated. The random-effects model as proposed by DerSimonian and Laird (50) was chosen *a priori* as the primary method to estimate all pooled estimates, appropriately supported by both clinical and statistical reasoning (51), as described previously (17). The extent and impact of between-study heterogeneity was assessed by inspecting the forest plots and by calculating τ^2 and the I^2 statistics, respectively. The 95% CIs around I^2 were calculated according to the non-central χ^2 approximation of Q (52). For meta-analyses with ≥ 3 trials, 95% Prediction Intervals (PI) (53, 54) were calculated to quantify of treatment effects of FFAs in a future trial.

All *P* values were 2-sided with $\alpha = 5\%$, except for the test of between-studies or between-subgroups heterogeneity ($\alpha = 10\%$) (55).

Additional analyses

Possible sources of heterogeneity in meta-analyses with 5 or more studies were sought through pre-specified mixed-effects subgroup analyses and random-effects meta-regression with the Knapp-Hartung adjustment (56).

Robustness of the results was *a priori* to be checked with sensitivity analyses based on (a) the duration of the FFA treatment, (b) the method error of the cephalometric analysis (where reported), and (c) the improvement of the GRADE classification.

Results

Study selection

From the initially identified 9115 records, 6342 remained after exclusion of duplicates and 6087 additional records were excluded on the basis of screening (Table 2). A total of 255 full texts were assessed for eligibility (Figure 1), with 10 articles (57–66) having at least one treatment arm with a FFA and being included in the systematic review. Two articles reported (59, 60) data from the same

study/cohort and were grouped together. Thus, 9 unique datasets were finally included in the qualitative and quantitative synthesis of this systematic review. In total, 2 authors were contacted; one of them responded, while the other one did not respond (communication details available upon request). The kappa scores before reconciliation for the selection, data extraction, and risk of bias assessment procedures were 0.855, 0.923, and 0.891, respectively (with asymptotic standard errors 0.103, 0.088, and 0.97), indicating almost perfect agreement.

Study characteristics and risk of bias within studies

The characteristics of the 10 included studies (9 datasets) are presented in Table 3. Seven of them took place at a university, one at a military academy, and one at a private practice, including a total of 418 subjects with a mean age of 13.2 years. The majority of the patients were treated with the original design of the corresponding FFA, while in two studies the FFAs were either modified or incorporated additional elements for maxillary expansion. All included studies provided data on skeletal and dentoalveolar changes, while five reported additionally on soft tissue cephalometric outcomes.

According to the Cochrane Collaboration's tool, the only one RCT identified (58) was judged to be in high risk of bias (Supplementary table 3a). The eight pCCTs (57, 59–66) scored an average of 22.4 points on the modified Downs and Black tool, which are actually not much, while none of them was considered to present serious methodological limitations (Supplementary table 3b).

Results of individual studies, synthesis of results and risk of bias across studies

A) Effectiveness of FFA treatment directly after appliance removal

Meta-analyses could be performed regarding only the short-term effectiveness of FFAs (i.e. from the time point of placement of the corresponding FFAs until immediately after their removal) compared to natural growth (as indicated by the data of the untreated control individuals) for 24 cephalometric variables, including eleven skeletal (five sagittal and six vertical), six dental, four soft tissue variables and three ratios (Table 4). In short, many skeletal, dental, and soft tissue variables were found to be significantly affected by FFA treatment.

With regard to the skeletal changes in the sagittal plane, the skeletal growth of the mandible was slightly affected by FFAs, with the SNB angle being on average 0.87° per year greater than the untreated group (Figure 2). Further, a statistically significant slight restriction effect on the maxillary growth of about 0.83° per year was induced by FFAs. The effect of FFAs on the skeletal relationships of the maxilla to the mandible was favourable, with the ANB angle being on average 1.74° decreased annually ($P < 0.001$) compared to the untreated group, indicating a moderate improvement of the skeletal Class II jaw relationships. Finally, as far as the vertical skeletal relationships are concerned, no significant effects could be found, except for annual increases of the SN-ML and SN-OP angles by 0.48° and 10.09° per year, respectively). The later indicates a clinically significant effect on the inclination of the occlusal plane during mandibular advancement.

With regard to the dentoalveolar changes, treatment effects were evident on all variables corresponding to the upper and lower dental arches. Significant retroclination of the upper incisors was observed compared to the untreated group, as seen from the 1s-SN (-7.50° per year) and 1s-NA angle (-4.24° per year). Additionally, the lower incisors were significantly proclined, as seen from the 1i-ML (7.99° per year), 1i-NB (4.20° per year), and 1i-VL angle (19.78° per year). Consequently, a statistically significant decrease in the interincisal angle was also noted (-8.32° per year).

The influence of FFAs on the soft tissues was significant for almost all available outcomes, with the mentolabial angle providing the more evident change (14.99° per year). Further, the H-angle was slightly decreased (-1.95° per year), while the N´SnPg´ angle was slightly increased (2.01° per year) compared to untreated patients. Finally no significant changes were observed regarding the cephalometric ratios investigated.

Risk of bias across studies

The GRADE assessments for the main outcomes after appliance removal ranged from low to high (Table 5). The quality of clinical recommendations was upgraded due to the magnitude of treatment effects for the upper incisor retroclination (via the 1s-SN angle) and the proclination of the mandibular incisors (via the 1i-ML angle). All judgements made for the GRADE analysis of each outcome are presented in detail in Supplementary table 4.

Additional analyses

The results of the included studies varied considerably and heterogeneity influenced the results of FFAs according to various patient-related factors (Table 6). Skeletal correction and facial convexity (via the ANB and NA-APg angles respectively) were significantly associated with patient sex. Post-peak patients showed a greater dentoalveolar effect with a greater emphasis on SN-NL and 1i-ML angles, and a greater reduction in 1s-SN angle compared with patients at pre-peak and peak skeletal growth stage. Finally, the impact of growth pattern on the treatment outcome was planned to be assessed, yet was not possible due to insufficient data reported in the included trials.

With regard to the appliance-related factors, for mandibular sagittal growth (via the SNB angle) and skeletal Class II correction (via the ANB angle) no statistical differences were observed, whereas the Forsus™ Fatigue Resistant Device was associated with the greatest proclination of the mandibular incisors. In addition, stepwise mandibular advancement was associated with greater retroclination of the upper incisors and greater proclination of the lower incisors compared to single step advancement.

Due to the limited number of included studies, it was not possible for an evaluation for the existence of reporting bias (including publication bias) to be performed.

Sensitivity analysis on the basis of treatment duration (Supplementary table 5) indicated that the duration of FFA treatment was significantly associated with the inclination of the upper and the lower incisors (via the 1s-SN and 1i-ML angles, respectively). Furthermore, the sensitivity analysis based on

the design of the included studies (Supplementary table 6) found no statistically significant differences between pCCTs and RCTs for the investigated angular measurements. Sensitivity analysis on the basis of the reported method error for each outcome could not be performed, due to incomplete reporting from the included studies. Finally, sensitivity analyses on the basis of the GRADE quality of recommendations were not feasible, as the only reason for downgrading the quality of recommendations was the inclusion of pCCTs over RCTs.

B) Effectiveness of FFA treatment after retention

Due to inadequate number of identified studies, no meta-analyses could be performed concerning the changes induced by FFAs after the retention phase.

Discussion

Summary of evidence

This systematic review included data from 418 patients and 10 RCTs and pCCTs, which assessed angular cephalometric changes induced by Class II treatment with FFAs. With regard to the ANB angle, the results from the random-effects meta-analyses indicated that FFAs had a statistically significant contribution in the improvement of skeletal Class II relationship (as seen by the average annualized decrease of 1.74^{o} of the ANB angle of the treated patients in comparison to untreated controls). This improvement was accomplished with approximately equal contributions from mandibular growth augmentation (0.87^{o} per year) and restriction of maxillary growth (0.83^{o} per year). However, the skeletal contribution in the sagittal plane to the correction of Class II malocclusion can be considered clinically small, in concordance with previous studies (8, 14, 15). Furthermore, in agreement with previous findings (67–69), the FFA treatment was associated with a significant increase of inclination of the occlusal plane relative to anterior cranial base.

The effects of FFAs were more pronounced on the dentition, where the maxillary incisors were significantly retroclined and the mandibular incisors were significantly proclined, resulting, thus, in a greater interincisal angle, agreeing with previous evidence (13, 69–71). Contradictory results regarding the effect of FFAs on soft tissues have been reported (13, 23, 69, 72). Based on the results of this study, soft tissues were significantly affected in favour of profile improvement by FFAs. Finally, no significant changes were observed in the investigated cephalometric ratios.

Several patient-related factors might explain the considerable variation in the treatment outcome among the included studies. The impact of patient's sex on the respective treatment could not be formally assessed in this study with separate sub-populations of male and female patients, due to the incomplete reporting of data in the original articles. The results of the meta-regression with the male/female ratio from each study indicate that the Class II correction and the skeletal facial convexity (through the ANB and NA-APg angles respectively) directly after the appliance removal

might be different between boys and girls, but further research is required to confirm or refute this. However, in a retrospective study, Rizell *et al*., (73) failed to identify a statistically significant association between gender and outcome of functional treatment.

The investigation of the patients' growth pattern and its influence on the FFA treatment results was originally planned but was not possible, due to the lack of data.

When the skeletal growth stage is taken into account, outcomes of FFA treatment were found to differ between patients before or during and patients after the growth peak. According to a previous study, dentoalveolar effects seem to increase in post-peak patients (74). Moreover, skeletal effects seem to be more pronounced in patients treated before (74, 75) or during the growth peak (8). The results of the present study indicate a trend towards more favourable skeletal sagittal changes in the pre-peak/peak than in the post-peak growth stage, although no statistical inference could be confirmed.

With regard to appliance-related factors, considerable differences were found among the seven different FFAs used in the included studies. Among them, the Forsus™ Fatigue Resistant Device was associated with the greatest proclination of the lower incisors (via the 1i-ML angle). A comparison of the effectiveness of the included FFAs is presented in Supplementary figures 1-3 regarding three skeletal variables and in Supplementary figures 4 and 5 regarding two dentoalveolar variables. However, these comparisons should be interpreted with caution due to the presence of considerable heterogeneity and their indirect nature. In a comparative evaluation of the Forsus™ Nitinol Flat Spring and the Jasper Jumper (15), the superiority of the latter in advancing the mandible was reported. Finally, contrary to the present study, a previous systematic review (8) reported that the Herbst appliance was found to produce more favourable results in matters of increased mandibular growth compared to other functional appliances. However, only the Herbst and the Mandibular Anterior Repositioning Appliance appliance were included from FFAs, while no quantitative synthesis was conducted.

The three different designs of the Herbst appliance that were used in the included studies were compared separately, with some significant differences among them (Table 5). Burkhardt *et al.*, (76) compared acrylic-splint and stainless-steel crown Herbst and reported that both the appliances produced similar skeletal changes. According to a previous assessment of the efficacy of three different Herbst anchorage systems (77), greater lower incisor proclination was observed with the cast-splint design than in the banded designs, although none of the designs could prevent mandibular anchorage loss. In the last years, attempts have been made to prevent anterior anchorage loss during mandibular protraction by utilising miniscrew implants, without however, particular success (78).

Moreover, the comparison between stepwise activation and maximum mandibular advancement did not reveal any significant differences, with the exception of the greater maxillary incisor retroclination that was observed with the former. Contrary to these results, a controlled trial investigating the effects of the Herbst appliance (72), reported that stepwise mandibular advancement produced greater

skeletal changes compared to single step advancement, while the dentoalveolar changes were more pronounced when the activation was performed at a single step.

Strengths and limitations

The strengths of the present review include the pre-defined protocol, the thorough and unrestricted literature search, and the strict methodology that was carefully followed during every stage of it following specific and detailed guidelines (31–33, 46). In addition, the 10 included studies enabled adequately-powered meta-analyses for many important treatment outcomes, for which the required meta-analysis sample size was fulfilled almost exclusively. Since a random-effects model was used for data synthesis, the results of the present study provide the average of the FFA effects across the included studies. Heterogeneity was explained in most cases by the pre-defined subgroup analyses, while sensitivity analyses indicated that the results were fairly robust. In addition, data from a minimum of five studies was considered as adequate to perform a subgroup analysis, in order to minimize multiple testing. Although most of the included trials took place at university settings, their findings could possibly be generalized to the average patient, due to their broad inclusion criteria. Finally, the vast majority of the included trials (89%) were pCCTs, which represent more "realistic" situations in matters of the daily clinical practice, compared to RCTs, and thus possibly strengthening the applicability of the outcomes.

Nevertheless, although no serious methodological limitations were found in the included original studies, their quality could have been better. Moreover, none of the included studies provided results concerning the number and the experience of the respective clinicians, which could introduce proficiency bias (79). Unfortunately, treatment results concerning the long-term effects of FFAs were not reported in the selected studies, precluding an assessment of the results' stability in current evaluation. Finally, the limited number of the eligible studies prevented the investigation of all the originally planned patient- and appliance-related factors in this review, while reporting biases could not be formally assessed (80).

Conclusions

According to existing evidence, the following conclusions can be drawn on the short-term effectiveness of FFAs:

The treatment effects of FFAs on the skeletal tissues in patients with Class II malocclusion excluding the effects of normal growth, were small and probably, of minor clinical importance.

The treatment of Class II malocclusion with FFAs was associated with small stimulation of mandibular growth, small inhibition of maxillary growth, and with more pronounced dentoalveolar and soft tissue changes.

Patient- and appliance-related factors seem to influence the treatment outcomes, yet complementary research is required to thoroughly investigate the respective effects.

The long-term effects of FFAs could not be properly assessed, because of insufficient number of relative trials at present.

Taking into account the clinical recommendations derived from the GRADE framework, high GRADE assessments could be drawn regarding the 1s-SN and 1i-ML angles exclusively. Particularly:

Clinicians should confidently expect an average reduction in the 1s-SN angle of 7.50° per year with the use of FFAs compared to untreated patients with the use of FFAs.

Clinicians should confidently expect an average increase in the 1i-ML angle of 7.99° per year with the use of FFAs compared to untreated patients.

Recommendations concerning the effectiveness of FFA treatment on the restriction of maxillary growth, advancement of the mandible, correction of skeletal Class II malocclusion, mandibular plane, and nasolabial angles are weaker and future research could affect them.

Treatment of Class II malocclusion with FFAs seems to be not as effective as believed in matters of skeletal correction. Additional studies are required for a thorough assessment of the skeletal, dental, and soft tissue outcomes of FFAs in the long-term. The provision of detailed data from these studies regarding patients' characteristics (gender, growth pattern, skeletal maturation), particular features of the used functional appliance (the exact appliance design and possible incorporation of additional elements), as well as the followed retention scheme should be considered. Finally, in order to enable also the assessment of linear variables, the magnification factor of the lateral cephalometric radiographs should be reported in each of the respective trials.

References

1. Pancherz, H., Zieber, K. and Hoyer, B. (1997) Cephalometric characteristics of Class II division 1 and Class II division 2 malocclusions: a comparative study in children. *The Angle Orthodontist*, 67, 111–120.
2. Kim Y.H. (1979) A comparative cephalometric study of Class II, Division 1 nonextraction and extraction cases. *The Angle Orthodontist*, 49, 77–84.
3. McNamara, J.A. Jr. (1981) Components of class II malocclusion in children 8-10 years of age. *The Angle Orthodontist*, 51, 177–202.
4. Paulsen, H.U., Karle, A., Bakke, M. and Herskind, A. (1995) CT-scanning and radiographic analysis of temporomandibular joints and cephalometric analysis in a case of Herbst treatment in late puberty. *European Journal of Orthodontics,* 17, 165–175.
5. McNamara, J.A. Jr., Howe, R.P. and Dischinger, T.G. (1990) A comparison of the Herbst and Fränkel appliances in the treatment of Class II malocclusion. *American Journal of Orthodontics and Dentofacial Orthopedics*, 98, 134–144.
6. Franchi, L., Baccetti, T. and McNamara, J.A. Jr. (1999) Treatment and posttreatment effects of acrylic splint Herbst appliance therapy. *American Journal of Orthodontics and Dentofacial Orthopedics*, 115, 429–438.
7. Woodside, D.G., Metaxas, A. and Altuna, G. (1987) The influence of functional appliance therapy on glenoid fossa remodeling. *American Journal of Orthodontics and Dentofacial Orthopedics*, 92, 181–198.
8. Cozza, P., Baccetti, T., Franchi, L., De Toffol, L. and McNamara, J.A. Jr. (2006) Mandibular changes produced by functional appliances in Class II malocclusion: a systematic review. *American Journal of Orthodontics and Dentofacial Orthopedics*, 129, 599.e1–12.
9. Flores-Mir, C., Ayeh, A., Goswani, A. and Charkhandeh, S. (2007) Skeletal and dental changes in Class II division 1 malocclusions treated with splint-type Herbst appliances. A systematic review. *The Angle Orthodontist*, 77, 376–381.
10. Covell, D.A. Jr., Trammell, D.W., Boero, R.P. and West, R. (1999) A cephalometric study of class II Division 1 malocclusions treated with the Jasper Jumper appliance. *The Angle Orthodontist*, 69, 311–320.
11. Croft, R.S., Buschang, P.H., English, J.D. and Meyer, R. (1999) A cephalometric and tomographic evaluation of Herbst treatment in the mixed dentition. *American Journal of Orthodontics and Dentofacial Orthopedics*, 116, 435–443.
12. Hägglund, P., Segerdal, S. and Forsberg, C.M. (2008) The integrated Herbst appliance–treatment effects in a group of adolescent males with Class II malocclusions compared with growth changes in an untreated control group. *European Journal of Orthodontics*, 30, 120–127.

13. Küçükkeleş, N., Ilhan, I. and Orgun, I.A. (2007) Treatment efficiency in skeletal Class II patients treated with the jasper jumper. *The Angle Orthodontist*, 77, 449–456.
14. Lai, M. and McNamara, J.A. Jr. (1998) An evaluation of two-phase treatment with the Herbst appliance and preadjusted edgewise therapy. *Seminars in Orthodontics*, 4, 46–58.
15. Darda, M., Goel, S. and Gupta, R. (2010) A cephalometric comparison of the dentoskeletal changes in class II malocclusion by using Jasper Jumper and Forsus-A clinical study. *International Journal of Contemporary Dentistry*, 1.
16. Cope, J.B., Buschang, P.H., Cope, D.D., Parker, J. and Blackwood, H.O. 3rd. (1994) Quantitative evaluation of craniofacial changes with Jasper Jumper therapy. *The Angle Orthodontist*, 64, 113–122.
17. Koretsi, V., Zymperdikas, V.F., Papageorgiou, S.N. and Papadopoulos, M.A. (2014) Treatment effects of removable functional appliances in patients with Class II malocclusion: a systematic review and meta-analysis. *European Journal of Orthodontics*, November 13, (Epub ahead of print).
18. Schäfer, K., Ludwig, B., Meyer-Gutknecht, H. and Schott, T.C. (2015) Quantifying patient adherence during active orthodontic treatment with removable appliances using microelectronic wear-time documentation. *European Journal of Orthodontics,* 37, 73–80.
19. Sahm, G., Bartsch, A. and Witt, E. (1990) Micro-electronic monitoring of functional appliance wear. *European Journal of Orthodontics*, 12, 297–301.
20. Chen, J.Y., Will, L.A. and Niederman, R. (2002) Analysis of efficacy of functional appliances on mandibular growth. *American Journal of Orthodontics and Dentofacial Orthopedics*, 122, 470–476.
21. Flores-Mir, C. and Major, P.W. (2006) A systematic review of cephalometric facial soft tissue changes with the Activator and Bionator appliances in Class II division 1 subjects. *European Journal of Orthodontics*, 28, 586–593.
22. Flores-Mir, C. and Major, P.W. (2006) Cephalometric facial soft tissue changes with the twin block appliance in Class II division 1 malocclusion patients. A systematic review. *The Angle Orthodontist*, 76, 876–881.
23. Flores-Mir, C., Major, M.P. and Major, P.W. (2006) Soft tissue changes with fixed functional appliances in Class II division 1. *The Angle Orthodontist,* 76, 712–720.
24. Antonarakis, G.S. and Kiliaridis, S. (2007) Short-term anteroposterior treatment effects of functional appliances and extraoral traction on Class II malocclusion. A meta-analysis. *The Angle Orthodontist,* 77, 907–914.
25. Harrison, J.E., O'Brien, K.D. and Worthington, H.V. (2007) Orthodontic treatment for prominent upper front teeth in children. *Cochrane Database of Systematic Reviews*, 3, CD003452.

26. Barnett, G.A., Higgins, D.W., Major, P.W. and Flores-Mir, C. (2008) Immediate skeletal and dentoalveolar effects of the crown- or banded type Herbst appliance on Class II division 1 malocclusion. *The Angle Orthodontist*, 78, 361–369.
27. Bertl, M.H., Mandl, C. and Crismani, A.G. (2011) Do functional orthodontic appliances stimulate mandibular growth in class II division 1 patients? *International Journal of Stomatology & Occlusion Medicine*, 4, 45–53.
28. Niu, Y. and Zhou, H. (2011) Effect on functional appliances on mandibular growth on skeletal Class II malocclusion: a systematic review. *West China Journal of Stomatology*, 29, 384–388.
29. Perillo, L., Cannavale, R., Ferro, F., Franchi, L., Masucci, C., Chiodini, P. and Baccetti, T. (2011) Meta-analysis of skeletal mandibular changes during Frankel appliance treatment. *European Journal of Orthodontics*, 33, 84–92.
30. Perinetti, G., Primožič, J., Furlani, G., Franchi, L. and Contardo, L. (2014) Treatment effects of fixed functional appliances alone or in combination with multibracket appliances: A systematic review and meta-analysis. *The Angle Orthodontist,* September 4, (Epub ahead of print).
31. Higgins, J.P.T. and Green, S. (2011) Cochrane handbook for systematic reviews of interventions [version 5.1.0, updated March 2011]. *The Cochrane Collaboration*. Available at: www.cochrane-handbook.org.
32. Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J. and Moher, D. (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of Clinical Epidemiology*, 62, e1–e34.
33. Beller, E.M., Glasziou, P.P., Altman, D.G., Hopewell, S., Bastian, H., Chalmers, I., Gøtzsche, P.C., Lasserson, T., Tovey, D. and PRISMA for Abstracts Group. (2013) PRISMA for Abstracts: Reporting Systematic Reviews in Journal and Conference Abstracts. *PLoS Medicine*, 10, e1001419.
34. Adams, J.W. (1940) Correction of errors in cephalometric roentgenograms. *The Angle Orthodontist*, 10, 3–13.
35. Athanasiou, A.E. (1995) *Orthodontic Cephalometry*. Mosby-Wolfe, London, UK.
36. Gandini, P., Mancini, M. and Andreani, F. (2006) A comparison of handwrist bone and cervical vertebral analyses in measuring skeletal maturation. *The Angle Orthodontist*, 76, 984–989.
37. Rasool, G., Bashir, U. and Kundi, I. (2010) Comparative evaluation between cervical vertebrae and hand-wrist maturation for assessment of skeletal maturity orthodontic patients. *Pakistan Oral & Dental Journal*, 30, 85–95.
38. Alkhal, H.A., Wong, R.W. and Rabie, A.B. (2008) Correlation between chronological age, cervical vertebral maturation and Fishman's skeletal maturity indicators in Southern Chinese. *The Angle Orthodontist*, 78, 591–596.

39. Bala, M., Pathak, A. and Jain, R.L. (2010) Assessment of skeletal age using MP 3 and hand-wrist radiographs and its correlation with dental and chronological ages in children. *Journal of the Indian Society of Pedodontics and Preventive Dentistry*, 28, 95–99.
40. Downs, S.H. and Black, N. (1998) The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health*, 52, 377–384.
41. Harmelink, K. (2010) *Effects of neurolytic blocks (botulinum toxin type A and motor branch block) in patients with a Stiff Knee Gait: A systematic review*. Master thesis. Utrecht University.
42. Peters, J.L., Sutton, A.J., Jones, D.R., Abrams, K.R. and Rushton, L. (2008) Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology*, 61, 991–996.
43. Begg, C.B. and Mazumdar, M. (1994) Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 50, 1088–1101.
44. Egger, M., Smith, G.D., Schneider, M. and Minder, C. (1997) Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315, 629–634.
45. Duval, S. and Tweedie, R. (2000) Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56, 455–463.
46. Guyatt, G.H., Oxman, A.D., Schünemann, H.J., Tugwell, P. and Knottnerus, A. (2011) GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *Journal of Clinical Epidemiology*, 64, 380–382.
47. Sloan, J., Symonds, T., Vargas-Chanes, D. and Fridley, B. (2003) Practical guidelines for assessing the clinical significance of health-related quality of life changes within clinical trials. *Drug Information Journal*, 37, 23–31.
48. Morissette, K., Tricco, A.C., Horsley, T., Chen, M.H. and Moher, D. (2011) Blinded versus unblinded assessments of risk of bias in studies included in a systematic review. *The Cochrane Database of Systematic Reviews*, 9, MR000025.
49. Cohen, J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.
50. DerSimonian, R. and Laird, N. (1986) Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7, 177–188.
51. Papageorgiou, S.N. (2014) Meta-analysis for orthodontists: Part I–How to choose effect measure and statistical model. *Journal of Orthodontics*, 41, 317–326.
52. Orsini, N., Bottai, M., Higgins, J. and Buchan, I. (2006) Heterogi: Stata module to quantify heterogeneity in a meta-analysis. *Statistical Software Components*. Available at: www.EconPapers.repec.org/RePEc:boc:bocode:s449201.

53. Higgins, J.P., Thompson, S.G. and Spiegelhalter, D.J. (2009) A re-evaluation of random effects meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172, 137–159.
54. Graham, P.L. and Moran, J.L. (2012) Robust meta-analytic conclusions mandate the provision of prediction intervals in meta-analysis summaries. *Journal of Clinical Epidemiology*, 65, 503–510.
55. Ioannidis, J.P. (2008) Interpretation of tests of heterogeneity and bias in meta-analysis. *Journal of Evaluation in Clinical Practice*, 14, 951–957.
56. Knapp, G. and Hartung, J. (2003) Improved tests for a random-effects meta-regression with a single covariate. *Statistics in Medicine*, 22, 2693–2710.
57. Alali, O.H. (2014) A prospective controlled evaluation of Class II division 1 malocclusions treated with fixed lingual mandibular growth modificator. *The Angle Orthodontist*, 84, 527–533.
58. Baysal, A. and Uysal, T. (2011) Soft tissue effects of Twin Block and Herbst appliances in patients with Class II division 1 mandibular retrognathy. *European Journal of Orthodontics*, 35, 71–81.
59. de Almeida, M.R., Henriques, J.F., de Almeida, R.R., Weber, U. and McNamara, J.A. Jr. (2005) Short-term treatment effects produced by the Herbst appliance in the mixed dentition. *The Angle Orthodontist*, 75, 540–547.
60. de Almeida, M.R., Flores-Mir, C., Brandão, A.G., de Almeida, R.R. and de Almeida-Pedrin, R.R. (2008) Soft tissue changes produced by a banded-type Herbst appliance in late mixed dentition patients. *World Journal of Orthodontics*, 9, 121–131.
61. Gunay, E.A., Arun. T. and Nalbantgil, D. (2011) Evaluation of the Immediate Dentofacial Changes in Late Adolescent Patients Treated with the Forsus(™) FRD. *European Journal of Dentistry*, 5, 423–432.
62. Karacay, S., Akin, E., Olmez, H., Gurton, A.U. and Sagdic, D. (2006) Forsus Nitinol Flat Spring and Jasper Jumper corrections of Class II division 1 malocclusions. *The Angle Orthodontist*, 76, 666–672.
63. Latkauskienė, D. (2012) *Treatment of angle class II malocclusion with the crown herbst appliance*. Doctoral dissertation. Kauno medicinos universitetas.
64. Oztoprak, M.O., Nalbantgil, D., Uyanlar, A. and Arun, T. (2012) A cephalometric comparative study of class II correction with Sabbagh Universal Spring (SUS(2)) and Forsus FRD appliances. *European Journal of Dentistry*, 6, 302–310.
65. Phelan, A., Tarraf, N.E., Taylor, P., Hönscheid, R., Drescher, D., Baccetti, T. and Darendeliler, M.A. (2012) Skeletal and dental outcomes of a new magnetic functional appliance, the Sydney Magnoglide, in Class II correction. *American Journal of Orthodontics and Dentofacial Orthopedics*, 141, 759–772.
66. Uyanlar, A., Nalbantgil, D. and Arun, T. (2014) Evaluation of Dentofacial Changes Caused with Sabbagh Universal Spring in Mandibular Retrognathic Patients: A Pilot Study. *Dentistry*, 4, 237.

67. Alvares, J.C.D.C., Cançado, R.H., Valarelli, F.P., Freitas, K.M.S.D. and Angheben, C.Z. (2013) Class II malocclusion treatment with the Herbst appliance in patients after the growth peak. *Dental Press Journal of Orthodontics*, 18, 38–45.
68. Siara-Olds, N.J., Pangrazio-Kulbersh, V., Berger, J. and Bayirli, B. (2010) Long-term dentoskeletal changes with the Bionator, Herbst, Twin Block, and MARA functional appliances. *The Angle Orthodontist*, 80, 18–29.
69. Nalbantgil, D., Arun, T., Sayinsu, K. and Fulya, I. (2005) Skeletal, dental and soft-tissue changes induced by the Jasper Jumper appliance in late adolescence. *The Angle Orthodontist*, 75, 426–436.
70. Heinig, N. and Göz, G. (2001) Clinical application and effects of the Forsus spring. A study of a new Herbst hybrid. *Journal of Orofacial Orthopedics*, 62, 436–450.
71. Cacciatore, G., Ghislanzoni, L.T., Alvetro, L., Giuntini, V. and Franchi, L. (2014) Treatment and posttreatment effects induced by the Forsus appliance: A controlled clinical study. *The Angle Orthodontist*, 84, 1010–1017.
72. Purkayastha, S.K., Rabie, A.B. and Wong, R. (2008) Treatment of skeletal class II malocclusion in adults: stepwise vs single-step advancement with the Herbst appliance. *World Journal of Orthodontics*, 9, 233–243.
73. Rizell, S., Svensson, B., Tengström, C. and Kjellberg, H. (2006) Functional appliance treatment outcome and need for additional orthodontic treatment with fixed appliance. *Swedish Dental Journal*, 30, 61–68.
74. Frye, L., Diedrich, P.R. and Kinzinger, G.S. (2009) Class II treatment with fixed functional orthodontic appliances before and after the pubertal growth peak - a cephalometric study to evaluate differential therapeutic effects. *Journal of Orofacial Orthopedics*, 70, 511–527.
75. Aras, A., Ada, E., Saracoğlu, H., Gezer, N.S. and Aras, I. (2011) Comparison of treatments with the Forsus fatigue resistant device in relation to skeletal maturity: a cephalometric and magnetic resonance imaging study. *American Journal of Orthodontics and Dentofacial Orthopedics*, 140, 616–625.
76. Burkhardt, D.R., McNamara, J.A. Jr. and Baccetti, T. (2003) Maxillary molar distalization or mandibular enhancement: a cephalometric comparison of comprehensive orthodontic treatment including the pendulum and the Herbst appliances. *American Journal of Orthodontics and Dentofacial Orthopedics*, 123, 108–116.
77. Weschler, D. and Pancherz, H. (2005) Efficiency of three mandibular anchorage forms in Herbst treatment: a cephalometric investigation. *The Angle Orthodontist*, 75, 23–27.
78. Bremen, J.V., Ludwig, B. and Ruf, S. (2014) Anchorage loss due to Herbst mechanics-preventable through miniscrews? *European Journal of Orthodontics,* December 1, (Epub ahead of print).

79. Krishna, R., Maithreyi, R. and Surapaneni, K.M. (2010) Research bias: a review for medical students. *Journal of Clinical and Diagnostic Research*, 4, 2320–2324.
80. Papageorgiou, S.N. (2014) Meta-analysis for orthodontists: Part II–Is all that glitters gold? *Journal of Orthodontics*, 41, 327–336.

Figure legends

Figure 1. PRISMA flow diagram for the selection of studies.

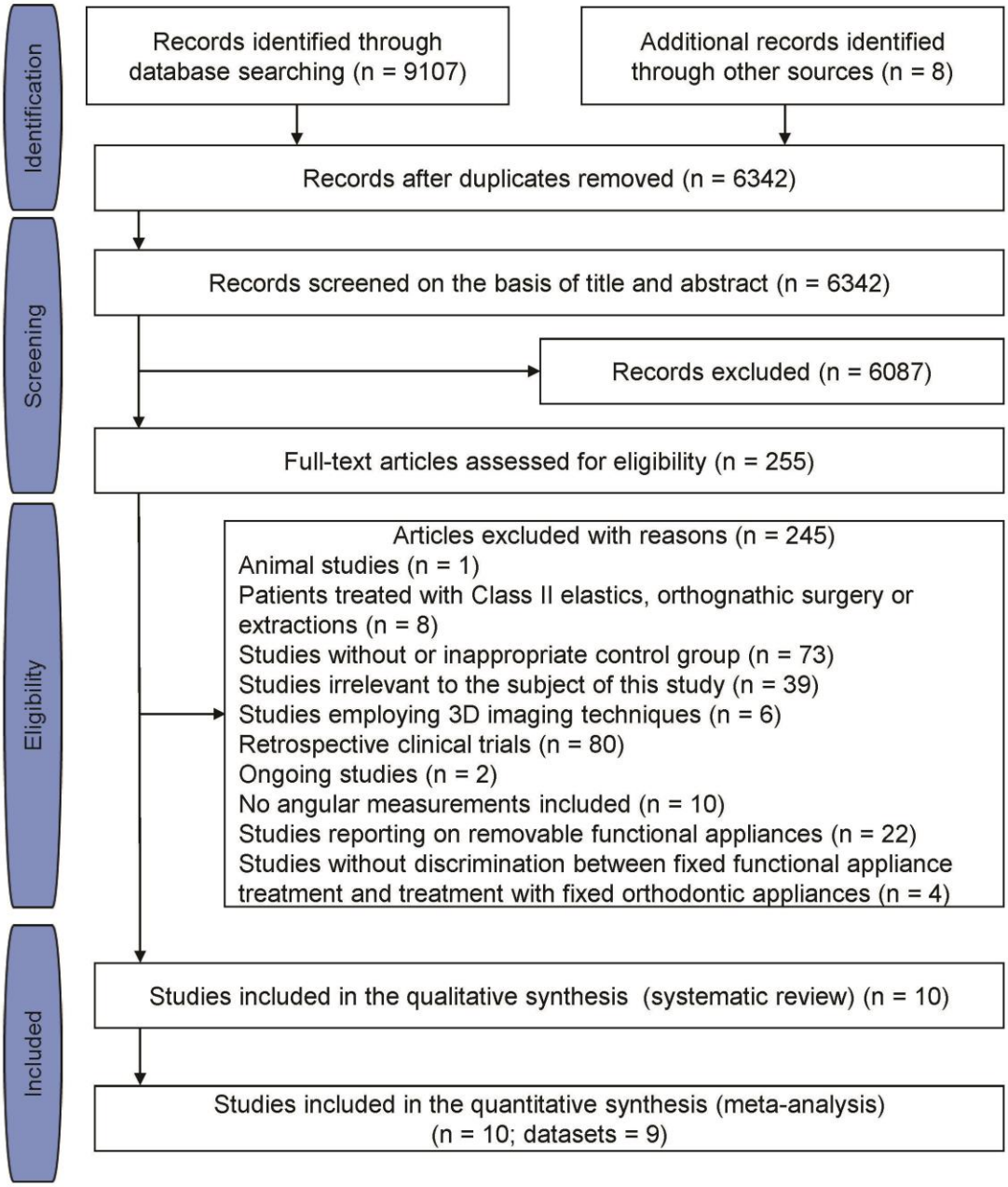
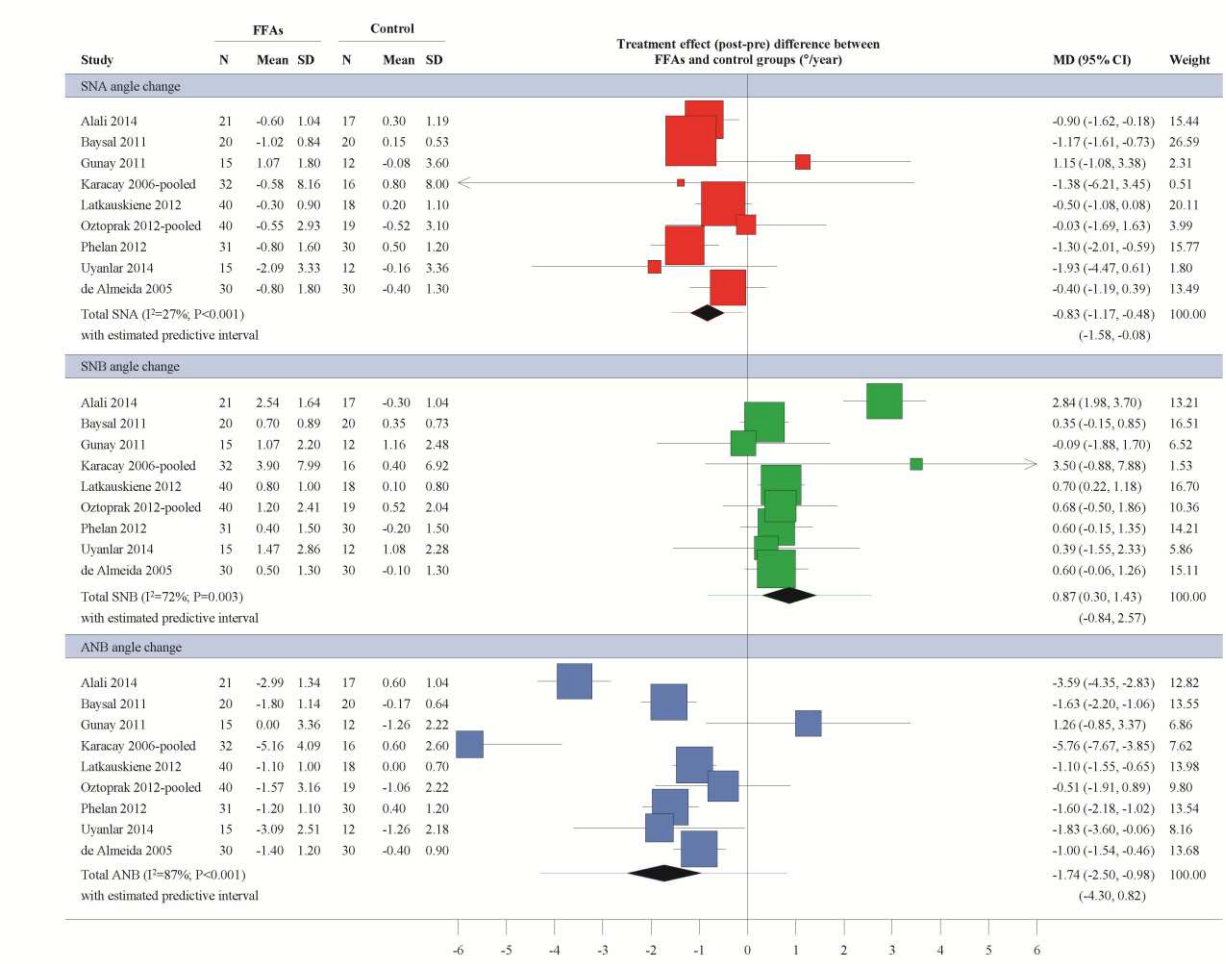


Figure 2. Forest plot of the mean difference of the SNA, SNB, and ANB angles (in °/year) between FFAs and control groups based on the random-effects model together with the 95% confidence interval and the 95% prediction interval. Studies on the right indicate that fixed functional treatment results in decreased SNA angle/increased SNB angle/decreased ANB angle compared to the normal growth, while studies on the left indicate increased SNA angle/decreased SNB angle/increased ANB angle compared to the normal growth.



Tables

Table 1. Eligibility criteria used for the study selection.

Category	*Inclusion criteria*	*Exclusion criteria*
Participant characteristics	Studies on human patients with Class II malocclusion of any age or sex	Patients with craniofacial syndromes and/or cleft lip palate Patients with temporomandibular joint disorders Animal studies
Intervention	Orthodontic treatment with fixed functional appliances	Patients with Class II malocclusion treated with extractions, Class I elastics, orthognathic surgery or removable functional appliances
Comparison	Untreated patients with Class II malocclusion matched for age and gender	Studies without an untreated Class II control group
Outcome	Studies providing angular skeletal, dentoalveolar and soft tissue cephalometric measurements from lateral cephalometric analysis	Studies providing only linear cephalometric measurements Electromyographic evaluation Evaluation employing 3D imaging techniques Cost benefit analysis
Study design	Randomized controlled clinical trials Prospective controlled clinical trials	Unsupported opinion of expert Editor' s choices Replies to the author / editor Interviews Commentaries Books' / Conferences' abstracts Summaries Cross-sectional surveys Case series without a control Case reports or reports of cases Case-control observational studies Cohort studies Retrospective clinical trials Narrative reviews* Systematic reviews* Meta-analyses*

*After checking the reference lists for relevant papers.

Table 2. Number of excluded studies with reasons.

Reason for exclusion	Excluded articles on the basis of title and abstract	Excluded articles on the basis of full-text
Animal studies	1	1
Patients treated with extractions, Class II elastics or orthognathic surgery	89	8
Studies without or inappropriate comparison group	144	73
Investigation not relevant to the subject of this study	5604	39
Evaluation employing 3D imaging techniques	7	6
Books' / Conferences' abstracts	72	–
Cross-sectional surveys	4	–
Case reports or reports of cases	62	–
Retrospective clinical trials	68	80
Narrative reviews	9	–
Systematic reviews	23	–
Meta-analyses	4	–
Ongoing studies	–	2
No angular cephalometric measurements	–	10
Removable functional appliances	–	22
No discrimination between treatment with fixed functional appliances and fixed orthodontic appliances	–	4
Sum	6087	245

.

Table 3. Characteristics of the studies included in the current systematic review. CCT†, prospective controlled trial with historical control; CCT, prospective controlled trial; RCT, randomized controlled trial; Gp1/2, treatment group; Ctr, control group; ME, maxillary expansion; M/F, males/females; m, months; NR, not reported.

A/A	Study	Design	Setting	Characteristics of patients	Interventions	No. of patients (M/F)	Age in years (SD)	Skeletal growth stage	Treatment time*** (m)	Outcomes	Risk of bias	Conflict of interest
1	Alali, (57)	CCT	University; Syria	Class II/1 malocclusion, overjet > 4 mm, ANB > 4º and APg/NL < 80º, SNB < 76º, pubertal growth spurt peak	Fixed lingual mandibular growth modificator	Gp: 21 (10/11) Ctr: 17 (7/10)	Gp: 13.2 (0.9) Ctr: 12.5 (2.1)	Peak	Gp: 18.2 Ctr: 17.9	Skeletal Dental	16	NR
2	Baysal and Uysal, (58)	RCT	University; Turkey	ANB > 4°, SNB < 78°, overjet ≥ 5 mm, SN-GoGn 32° ± 6°, crowding in dental arches ≤ 4 mm, bilateral Class II molar and canine relationship ≥ 3.5 mm	Cast splint Herbst (ME)	Gp: 20 (9/11) Ctr: 20 (11/9)	Gp: 12.7 (1.4) Ctr: 12.2 (1.5)	Pre-peak and peak	Gp: 15.8 Ctr: 15.6	Skeletal Dental Soft tissue	High risk	Internal
3	de Almeida *et al.*, (59) and de Almeida *et al.*, (60)	CCT†	University; Brazil	Class II/1 malocclusion with bilateral distal molar relationship > one-half cusp, presence of mandibular deciduous second molars, ANB ≥ 4.5°	Herbst (modified)	Gp: 30 15/15)* Ctr: 30 (15/15)	Gp: 9.8 (NR)** Ctr: 9.7 (NR)	Pre-peak and peak	Gp: 12.0 Ctr: 12.0	Skeletal Dental Soft tissue	23	NR
4	Gunay *et al.*, (61)	CCT†	University; Turkey	Skeletal and dental class II malocclusion due to retrognathic mandible, normal or low-angle growth pattern, postpeak growth period, no extracted or congenitally missing permanent teeth, minimum crowding in the lower dental arch	Forsus FRD	Gp: 15 (6/9) Ctr: 12 (3/9)	Gp: 15.0 (1.2) Ctr: 14.1 (1.4)	Post-peak	Gp: 5.3 Ctr: 6.0	Skeletal Dental Soft tisue	18	NR
5	Karacay *et al.*, (62)	CCT	Military Medical Academy; Turkey	Active growth period, normal or mildly prognathic maxilla, retrognathic mandible, horizontal or normal growth pattern, class II molar relationship, overjet < 7 mm, minimum crowding, permanent dentition	Gp1: Forsus nitinol flat spring Gp2: Jasper jumper	Gp1: 16 (9/7) Gp2: 16 (10/6) Ctr: 16 (NR)	Gp1: 13.6 (1.2) Gp2: 14.0 (1.9) Ctr: 13.8 (1.4)	NR	Gp1: 5.3 Gp2: 5.2 Ctr: 6.0	Skeletal Dental	22	NR
6	Latkauskienė, (63)	CCT†	University; Lithuania	≥ end-to-end Class II molar relationship bilaterally or more severe, permanent dentition, no active hard tissue lesions, no previous orthodontic treatment or tooth extractions, no bone level problems, no TMJ complaints, no tooth size, form, and number anomalies, no facial development	Stainless steel crown Herbst	Gp: 40 (20/20) Ctr: 18 (11/7)	Gp: 13.6 (1.3) Ctr: 13.9 (1.6)	Peak and post-peak	Gp: 12.0 Ctr: 12.0	Skeletal Dental	25	NR

				or mental syndroms, no pregnancy									
7	Oztoprak *et al.*, (64)	CCT†	University; Turkey	Class II malocclusion due to retrognathic mandible, SN-MP = 25°-35°, post-peak growth period, no extracted or congenitally missing permanent teeth, minimum crowding in the lower arch	Gp1: Sabbagh universal spring Gp2: Forsus FRD	Gp1: 20 (9/11) Gp2: 20 (8/12) Ctr: 19 (5/14)	Gp1: 15.3 (1.2) Gp2: 15.1 (1.0) Ctr: 14.8 (1.3)	Post-peak	Gp1: 5.2 Gp2: 5.2 Ctr: 6.0	Skeletal Dental Soft tissue	23	NR
8	Phelan *et al.*, (65)	CCT†	Private practice; Germany	Class II/1 malocclusion of a half or full cusp, overjet ≥ 6 mm, ANB > 3.5°, and nonextraction treatment plan	Sydney magnoglide	Gp: 31 (19/12) Ctr: 30 (15/15)	Gp: 13.5 (1.2) Ctr: 13.0 (1.6)	Pre-peak, peak, and post-peak	Gp: 12.0 Ctr: 12.0	Skeletal Dental	22	External; non profit
9	Uyanlar *et al.*, (66)	CCT	University; Turkey	SNB < 80°, SN-ML ≤ 32°, post-peak growth period, no extracted or congenitally missing permanent teeth, minimum crowding in the lower dental arch	Sabbagh universal spring	Gp: 15 (7/8) Ctr: 12 (3/9)	Gp: 15.2 (1.1) Ctr: 14.1 (1.4)	Post-peak	Gp: 5.2 Ctr: 6.0	Skeletal Dental Soft tissue	19	NR

*Only for soft tissue measurements Gp: 29 (15/14) and Ctr: 28 (14/14).
**Only for soft tissue measurements Gp: 9.9 (NR).
***Control group received no intervention; it refers to the observation time.

Table 4. Details of the performed meta-analyses with tests on heterogeneity. FA, functional appliance group; Ctr, control group; MD, mean difference; CI, confidence interval; PI, prediction interval; NA, not applicable.

					Effect Size				Heterogeneity		
No.	Variable	Studies	Post-Pre in FA*	Post-Pre in Ctr*	MD	95% CI	P-Value	95% PI	P-Value	τ^2	I^2 (95% CI)
1	SNA	9**	↓(-0.57)	↑(0.13)	-0.83	-1.17,-0.48	**<0.001**	-1.58,-0.08	0.202	0.069	27% (0%,66%)
2	SNB	9**	↑(1.07)	↑(0.11)	0.87	0.30,1.43	**0.003**	-0.84,2.57	<0.001	0.435	72% (32%,84%)
3	SN-Pg***	2	↑(1.45)	↑(0.24)	1.29	-0.52,3.09	0.162	NA	<0.001	1.588	94% (NA)
4	ANB****	9**	↓(-1.95)	↓(-0.13)	-1.74	-2.50,-0.98	**<0.001**	-4.30,0.82	<0.001	1.019	87% (78%,92%)
5	NA-Apg****	5**	↓(-3.01)	↓(-0.96)	-1.86	-5.06,1.34	0.254	-13.68,9.95	<0.001	11.120	90% (78%,94%)
6	SGo:Nme (%)	4**	↑(1.31)	↑(1.29)	0.71	-0.09,1.51	0.080	-1.04,2.46	0.424	0.000	0% (0%,68%)
7	SN-ML	8**	↑(0.16)	↓(-0.43)	0.48	0.04,0.92	**0.031**	-0.07,1.03	0.949	0.000	0% (0%,56%)
8	NL-ML	2**	↑(0.28)	↑(0.15)	0.13	-0.95,1.20	0.818	NA	0.776	0.000	0% (NA)
9	SN-NL	6**	↑(0.41)	↓(-0.29)	0.54	-0.23,1.32	0.168	-1.30,2.39	0.181	0.284	34% (0%,73%)
10	SN-OP	4**	↑(8.44)	↓(-1.58)	10.09	7.22,12.96	**<0.001**	-0.96,21.13	0.099	4.442	52% (0%,82%)
11	y axis***	2**	↑(3.65)	↑(0.34)	3.06	-2.18,8.29	0.252	NA	0.002	12.881	90% (NA)
12	1s-SN****	6**	↓(-6.49)	↑(0.93)	-7.50	-10.88,-4.11	**<0.001**	-18.32,3.33	<0.001	12.209	79% (44%,89%)
13	1i-ML****	9**	↑(8.20)	↑(0.25)	7.99	3.56,12.42	**<0.001**	-8.34,24.32	<0.001	42.595	96% (95%,97%)
14	1s-1i*****	3**	↓(-4.59)	↓(-1.48)	-8.32	-13.38,-3.25	**0.001**	-57.04,40.41	0.190	8.038	40% (0%,82%)
15	1s-NA	2	↓(-3.38)	↑(1.16)	-4.24	-6.09,-2.40	**<0.001**	NA	0.233	0.572	30% (NA)
16	1i-NB	2	↑(4.60)	↑(0.36)	4.20	2.48,5.91	**<0.001**	NA	0.195	0.684	40% (NA)
17	1i-VL	3**	↑(18.67)	↓(-1.13)	19.78	15.50,24.06	**<0.001**	-24.38,63.94	0.129	7.319	51% (0%,85%)
18	N´SnPg´	2	↑(1.88)	↓(-0.19)	2.01	1.05,2.96	**<0.001**	NA	0.622	0.000	0% (NA)
19	Nasolabial angle	5**	↑(0.49)	↑(0.49)	0.03	-2.39,2.45	0.979	-4.83,4.89	0.347	0.802	10% (0%,68%)
20	Mentolabial angle	2	↑(10.00)	↓(-5.19)	14.99	8.09,21.88	**<0.001**	NA	0.233	7.337	30% (NA)
21	H angle	4**	↓(-2.14)	↓(-2.01)	-1.95	-3.16,-0.74	**0.002**	-4.61,0.70	0.908	0.000	0% (0%,68%)
22	ANSMe:Nme******	3**	↑(0.89)	↑(0.63)	*omitted*						
23	Gonial Ratio	3**	↑(1.99)	↑(0.23)	1.62	-0.66,3.90	0.164	-23.17,26.41	0.080	2.452	60% (0%,87%)
24	S-Ar/Ar-Go	2**	↓(-3.86)	↓(-4.77)	0.49	-3.48,4.46	0.809	NA	0.667	0.000	0% (NA)

*Results from random-effects meta-analysis of the post-pre differences in each group to provide an overview of the effect's direction.
**Pooled trial arms included.
***High heterogeneity identified; however, our confidence in the calculation of heterogeneity is limited, due to the small number of studies. Furthermore, it would affect only the estimation of the effect magnitude; not its direction (i.e. all studies lie on the same side of the forest plot).

****High heterogeneity identified; however, heterogeneity is explained by differences between subgroups. Caution is warranted on the interpretation of the overall effect estimate; estimates for subgroups are to be preferred.

*****Initial analysis included four studies (MD=-3.14; 95% CI=-15.95 to 9.67; P=0.631; τ^2=160.559; I^2=95%), but the study Alali, (57) (fixed lingual mandibular growth modificator) was omitted to achieve homogeneity.

******High heterogeneity identified, which remained unexplained; meta-analysis of three studies (MD=0.13; 95% CI=-2.17,2.42; P=0.915; τ^2=3.472; I^2=85%) omitted, as studies were distributed on both sides of the forest plot and elimination of a single study was not straightforward.

Table 5. GRADE summary of findings table for the main outcomes of the systematic review directly after treatment with fixed functional appliances. CI, confidence interval; Ctr, untreated control group; mo, month; FFA, fixed functional appliance.

Patients: receiving orthodontic treatment to improve Class II malocclusion
Settings: university clinics (Brazil, Lithuania, Syria, Turkey), private practice (Germany), and military academy (Turkey)
Intervention: FFAs (Fixed Lingual Mandibular Growth Modificator, Forsus Fatigue Resistant Device, Forsus Nitinol Flat Spring, Herbst, Jasper Jumper, Sabbagh Universal Spring, Sydney Magnoglide)
Comparison: untreated patients from follow-up or historical controls

	Illustrative comparative risks (95% CI)				
Outcomes	Assumed risk Untreated (Ctr) patients	Corresponding risk FFA patients	No. of patients (trials)	Quality of evidence (GRADE)	Comments
Annualized SNA change from baseline (follow-up: 5.2–15.8 mos)*	The SNA increased on average by 0.13° per year in the Ctr groups (range -0.52° to 0.80°)	The mean SNA decreased in the FFA groups by 0.83° per year (95% CI: 0.48°–1.17° decrease) compared to the Ctr groups	418 (9)	⊕⊕⊖⊖ Low	–
Annualized SNB change from baseline (follow-up: 5.2–15.8 mos)*	The SNB increased on average by 0.11° per year in the Ctr groups (range -0.30° to 1.16°)	The mean SNB increased in the FFA groups by 0.87° per year (95% CI: 0.30°–1.43° increase) compared to the Ctr groups	418 (9)	⊕⊕⊖⊖ Low	Effect magnitude affected by appliance
Annualized ANB change from baseline (follow-up: 5.2–15.8 mos)*	The ANB decreased on average by 0.13° per year in the Ctr groups (range -1.26° to 0.60°)	The mean ANB decreased in the FFA groups by 1.74° per year (95% CI: 0.98°–2.50° decrease) compared to the Ctr groups	418 (9)	⊕⊕⊖⊖ Low	Effect magnitude affected by (i) patient's gender and (ii) appliance
Annualized SN-ML change from baseline (follow-up: 5.2–15.8 mos)*	The SN-ML decreased on average by 0.43° per year in the Ctr groups (range -1.92° to 0.30°)	The mean SN-ML increased in the FFA groups by 0.48° per year (95% CI: 0.04°–0.92° increase) compared to the Ctr groups	360 (8)	⊕⊕⊖⊖ Low	–
Annualized 1s-SN change from baseline (follow-up: 5.2–15.8 mos)*	The 1s-SN increased on average by 0.93° per year in the Ctr groups (range 0.40° to 3.06°)	The mean 1s-SN decreased in the FFA groups by 7.50° per year (95% CI: 4.11°–10.88° decrease) compared to the Ctr groups	262 (6)	⊕⊕⊕⊕ Very high**	Effect magnitude affected by (i) patient's skeletal growth stage and (ii) construction bite
Annualized 1i-ML change from baseline (follow-up: 5.2–15.8 mos)*	The 1i-ML increased on average by 0.25° per year in the Ctr groups (range -0.80° to 1.00°)	The mean 1i-ML increased in the FFA groups by 7.99° per year (95% CI: 3.56°–12.42° increase) compared to the Ctr groups	418 (9)	⊕⊕⊕⊕ Very high**	Effect magnitude affected by (i) patient's skeletal growth stage and (ii) appliance
Annualized nasolabial angle change from baseline (follow-up: 5.2–15.8 mos)*	The nasolabial angle increased on average by 0.49° per year in the Ctr groups (range -1.34° to 1.81°)	The mean nasolabial angle increased in the FFA groups by 0.03° per year (95% CI: 2.39° decrease–2.45° increase) compared to the Ctr groups	210 (5)	⊕⊕⊖⊖ Low	–

All judgements start from "low", due to the vast inclusion of non-randomized studies.
*From cephalometric analysis.
**Upgraded by two for effect magnitude; very large effect (cephalometric norm + 2 SDs + 1° for method error), which was included in the mean effect, the confidence interval and the prediction interval, while no serious limitations were found. Furthermore, magnitude of incisor inclination change significantly associated with duration of functional appliance treatment.

Table 6. Details of the performed subgroup analyses. MD, mean difference; CI, confidence interval; P_{SG}, P values for difference between subgroups; FFRD, Forsus Fatigue Resistant Device; FLMGM, Fixed Lingual Mandibular Growth Modificator; FNFS, Forsus Nitinol Flat Spring; JJ, Jasper Jumper; SUS, Sabbagh Universal Spring; SM, Sydney Magnoglide; SS, stainless steel.

	SNA n=11			SNB n=11			ANB n=11			NA-APg n=6			SN-ML n=10		
	n	MD (95% CI)	P_{SG}	n	MD (95% CI)	P_{SG}	n	MD (95% CI)	P_{SG}	n	MD (95% CI)	P_{SG}	n	MD (95% CI)	P_{SG}
Patient-related															
Gender ratio	11		0.190	11		0.710	11		**0.080**	6		**0.002**	10		0.368
Skeletal growth stage															
Pre-peak and peak	3	-0.92 (-1.35,-0.50)	0.316	3	1.23 (-0.12,2.57)	0.401	3	-2.05 (-3.43,-0.67)	0.218	2	-4.89 (-9.21,-0.57)	0.135	3	0.46 (-0.09,1.00)	0.440
Post-peak	4	-0.11 (-1.81,1.59)		4	0.48 (-0.31,1.27)		4	-0.49 (-2.03,1.06)		4	0.72 (-3.06,4.50)		4	0.95 (-0.07,1.98)	
Appliance-related															
Appliance															
FFRD	2	1.31 (0.00,2.62)	0.284	2	0.43 (-0.65,1.49)	**0.082**	2	0.94 (-0.42,2.30)	**0.017**	2	4.08 (1.38,6.78)	**0.060**	2	1.41 (-0.39,3.21)	0.881
FLMGM	1	-0.90 (-1.62,-0.18)		1	2.84 (1.98,3.70)		1	-3.59 (-4.35,-2.83)		1	-7.16 (-8.93,-5.39)		1	0.30 (-0.92,1.52)	
FNFS	1	-1.48 (-7.22,4.26)		1	4.37 (-1.02,9.76)		1	-7.19 (-9.67,-4.71)					1	-0.43 (-9.76,8.90)	
Herbst	3	-0.75 (-1.27,-0.23)		3	0.55 (0.24,0.86)		3	-1.22 (-1.58,-0.86)		1	-2.75 (-3.79,-1.71)		2	0.50 (-0.12,1.11)	
JJ	1	-1.27 (-6.82,4.28)		1	2.62 (-2.40,7.64)		1	-4.32 (-6.41,-2.23)					1	-1.83 (-9.22,5.56)	
SUS	2	-1.64 (-3.18,-0.09)		2	0.55 (-0.63,1.72)		2	-1.77 (-2.90,-0.64)		2	-2.54 (-5.13,0.05)		2	0.74 (-0.51,1.98)	
SM	1	-1.30 (-2.01,-0.59)		1	0.60 (-0.15,1.35)		1	-1.60 (-2.18,-1.02)					1	0.20 (-0.75,1.15)	
Herbst design															
SS crown	1	-0.50 (-1.08,0.08)	**0.095 ***	1	0.70 (0.22,1.18)	0.607 *	1	-1.10 (-1.55,-0.65)	0.236 *						
Cast splint	1	-1.17 (-1.61,-0.74)		1	0.35 (-0.15,0.85)		1	-1.63 (-2.20,-1.06)					1	0.56 (-0.23,1.35)	0.802 *
Modified	1	-0.40 (-1.20,0.40)		1	0.60 (-0.06,1.26)		1	-1.00 (-1.54,-0.46)					1	0.40 (-0.57,1.37)	
Construction bite															
Single step	2	-0.86 (-1.60,-0.12)	0.444	2	0.44 (0.04,0.84)	0.452	2	-1.31 (-1.92,-0.69)	0.765	1	-2.75 (-3.79,-1.71)	0.436	2	0.50 (-0.12,1.11)	0.542
Stepwise	7	-0.27 (-1.20,0.66)		7	0.68 (0.27,1.09)		7	-1.89 (-3.37,-0.40)		4	0.72 (-3.06,4.50)		6	0.89 (-0.12,1.89)	

Table 6. *(continued)*

	SN-NL n=8			1s-SN n=8			1i-ML n=11			Nasolabial angle n=6		
	n	MD (95% CI)	P_{SG}	n	MD (95% CI)	P_{SG}	n	MD (95% CI)	P_{SG}	n	MD (95% CI)	P_{SG}
Patient-related												
Gender ratio	8		0.168	8		0.497	11		0.377	6		0.104
Skeletal growth stage												
Pre-peak and peak	2	0.06 (-0.47,0.59)	**0.052**	1	-4.24 (-6.19,-2.30)	**0.035**	3	0.16 (-5.12,5.43)	**0.010**	2	0.51 (-5.05,6.07)	0.824
Post-peak	4	1.93 (0.70,3.17)		4	-10.92 (-13.92,-7.92)		4	18.89 (12.50,25.28)		4	-0.44 (-3.64,2.76)	
Appliance-related												
Appliance												
FFRD	2	1.60 (0.08,3.11)	0.397	2	-10.60 (-15.89,-5.30)	0.151	2	24.51 (21.59,27.43)	**0.002**	2	-2.75 (-6.45,0.96)	0.348
FLMGM	1	0.00 (-0.63, 0.63)					1	-6.72 (-9.88,-3.56)				
FNFS	1	-1.79 (-6.55,2.97)		1	-11.54 (-17.93,-5.15)		1	11.03 (5.47,16.59)				
Herbst	1	0.20 (-0.78,1.18)		1	-4.24 (-6.19,-2.30)		3	3.92 (2.25,5.59)		2	0.51 (-5.05,6.07)	
JJ	1	0.03 (-4.21,4.27)		1	-12.03 (-20.39,-3.67)		2	16.15 (6.07,26.23)				
SUS	2	2.59 (0.46,4.72)		2	-11.07 (-14.71,-7.44)		2	13.31 (9.95,16.67)		2	3.01 (-1.71,7.73)	
SM				1	-2.20 (-4.24,-0.16)		1	0.40 (-1.05,1.85)				
Herbst design												
SS crown							1	5.40 (3.38,7.42)	0.102 *			**0.084 ***
cast splint							1	2.74 (1.33,4.15)		1	-2.15 (-6.26,1.96)	
modified							1	4.00 (1.58,6.42)		1	3.54 (-1.45,8.53)	
Construction bite												
Single step	1	0.20 (-0.78,1.18)	0.135	1	-4.24 (-6.19,-2.30)	**0.009**	2	3.06 (1.84,4.28)	**0.057**	2	0.51 (-5.05,6.07)	0.824
Stepwise	6	1.57 (0.42,2.72)		6	-11.13 (-13.71,-8.55)		7	15.34 (8.39,22.30)		4	-0.44 (-3.64,2.76)	

Bold values indicate statistically significant differences between subgroups.
*Mixed-effects subgroup analysis not possible, due to the small sample. Conventional inverse variance subgroup analysis reported.

Supplementary material

Supplementary table 1 Electronic databases searched, search strategies used and corresponding results.

Electronic database	*Search strategy used*	*Limits*	*Hits*
MEDLINE Searched via PubMed (1950 - week 1, October 2014) http://www.ncbi.nlm.nih.gov/pubmed/advanced	((((((maxill* AND (excess* OR prognath*)) OR (mandib* AND (deficien* OR retrognath* OR reposition* OR enhanc* OR advanc*)) OR functional OR orthopaedic* OR orthopedic* OR growth)) OR (Herbst OR "Magnetic telescopic device" OR "Ventral telescope" OR "Mandibular advancing repositioning splint" OR "Mandibular corrector appliance" OR "Biopedic appliance" OR "Ritto appliance" OR "Mandibular protraction appliance" OR "Mandibular anterior repositioning appliance" OR "MARA" OR "Functional mandibular advancer" OR "Jasper jumper" OR "Scandee tubular jumper" OR "Flex developer" OR "Adjustable bite corrector" OR "Bite fixer" OR "Forsus nitinol flat spring" OR "Forsus device" OR "Forsus appliance" OR "Twin force bite corrector" OR "Eureka spring" OR "Sabbagh spring" OR Activator OR Bionator OR "Bimler appliance" OR "Fraenkel appliance" OR "Frankel appliance" OR "Bass appliance" OR "Harvold appliance" OR "Andresen appliance" OR "Teuscher appliance" OR "Stoeckli appliance" OR "Stockli appliance" OR Biobloc OR "Bite jumper" OR "Bite jumping" OR "SII appliance" OR "Twin block")) AND ("class ii malocclusion" OR class ii div* OR class ii/* OR ("class ii" AND orthodont*)))) AND ((randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized controlled trials[mh] OR random allocation[mh] OR double-blind method[mh] OR single-blind method[mh] OR clinical trial[pt] OR clinical trials[mh]) OR ("clinical trial"[tw]) OR ((singl*[tw] OR doubl*[tw] OR trebl*[tw] OR tripl*[tw]) AND (mask*[tw] OR blind*[tw])) OR (placebos[mh] OR placebo*[tw] OR random*[tw] OR research design[mh:noexp] OR comparative study OR evaluation studies OR follow-up studies[mh] OR prospective studies[mh] OR control*[tw] OR prospectiv*[tw] OR volunteer*[tw]))	*No limitations*	**1493**
EMBASE Searched via embase biomedical answers (<1966 – week 1, October 2013) http://www.embase.com/search/advanced	(maxill* AND (excess* OR prognath*)) OR (mandib* AND (deficien* OR retrognath* OR reposition* OR enhanc* OR advanc*)) OR functional OR orthopaedic* OR orthopedic* OR 'growth'/exp OR herbst OR 'magnetic telescopic device' OR 'ventral telescope' OR 'mandibular advancing repositioning splint' OR 'mandibular corrector appliance' OR 'biopedic appliance' OR 'ritto appliance' OR 'mandibular protraction appliance' OR 'mandibular anterior repositioning appliance' OR 'mara' OR 'functional mandibular advancer' OR 'jasper jumper' OR 'scandee tubular jumper' OR 'flex developer' OR 'adjustable bite corrector' OR 'bite fixer' OR 'forsus nitinol flat spring' OR 'forsus device' OR 'forsus appliance' OR 'twin force bite corrector' OR 'eureka spring' OR 'sabbagh spring' OR activator OR bionator OR 'bimler appliance' OR 'fraenkel appliance' OR 'frankel appliance' OR 'bass appliance' OR 'harvold appliance' OR 'andresen appliance' OR 'teuscher appliance' OR 'stoeckli appliance' OR 'stockli appliance' OR biobloc OR 'bite jumper' OR 'bite jumping' OR 'sii appliance' OR 'twin block' AND ('class ii malocclusion' OR class AND ii AND div* OR class AND ii OR ('class ii' AND orthodont*)) AND ('clinical trial'/exp OR 'comparative study'/exp OR 'controlled clinical trial'/exp OR 'double blind procedure'/exp OR 'prospective study'/exp OR 'randomized controlled trial'/exp) AND 'malocclusion'/exp	*Limitations: Study type: clinical trial, comparative study, controlled clinical trial, double blind procedure, prospective study, randomized controlled trial, Disease: malocclusion*	**549**
Cochrane Database of Systematic Reviews Searched via The Cochrane Library on October 10, 2014 http://onlinelibrary.wiley.com/o/cochrane/cochrane_search_fs.html?newSearch=true	(maxill* AND (excess* OR prognath*)) OR (mandib* AND (deficien* OR retrognath* OR reposition* OR enhanc* OR advanc*)) OR functional OR orthopaedic* OR orthopedic* OR growth OR (Herbst OR "Magnetic telescopic device" OR "Ventral telescope" OR "Mandibular advancing repositioning splint" OR "Mandibular corrector appliance" OR "Biopedic appliance" OR "Ritto appliance" OR "Mandibular protraction appliance" OR "Mandibular anterior repositioning appliance" OR "MARA" OR "Functional mandibular advancer" OR "Jasper jumper" OR "Scandee tubular jumper" OR "Flex developer" OR "Adjustable bite corrector" OR "Bite fixer" OR "Forsus nitinol flat spring" OR "Forsus device" OR "Forsus appliance" OR "Twin force bite corrector" OR "Eureka spring" OR "Sabbagh spring" OR Activator OR Bionator OR "Bimler appliance" OR "Fraenkel appliance" OR "Frankel appliance" OR "Bass appliance" OR "Harvold appliance" OR "Andresen appliance" OR "Teuscher appliance" OR "Stoeckli appliance" OR "Stockli appliance" OR Biobloc OR "Bite jumper" OR "Bite jumping" OR "SII appliance" OR "Twin block") AND ("class ii malocclusion" OR class ii div* OR class ii/ OR ("class ii" AND orthodont*))	*No limitations*	**628**
Cochrane Central Register of Controlled Trials Searched via The Cochrane Library on October 10, 2014 http://onlinelibrary.wiley.com/o/cochrane/cochrane_search_fs.html?newSearch=true	(maxill* AND (excess* OR prognath*)) OR (mandib* AND (deficien* OR retrognath* OR reposition* OR enhanc* OR advanc*)) OR functional OR orthopaedic* OR orthopedic* OR growth OR (Herbst OR "Magnetic telescopic device" OR "Ventral telescope" OR "Mandibular advancing repositioning splint" OR "Mandibular corrector appliance" OR "Biopedic appliance" OR "Ritto appliance" OR "Mandibular protraction appliance" OR "Mandibular anterior repositioning appliance" OR "MARA" OR "Functional mandibular advancer" OR "Jasper jumper" OR "Scandee tubular jumper" OR "Flex developer" OR "Adjustable bite corrector" OR "Bite fixer" OR "Forsus nitinol flat spring" OR "Forsus device" OR "Forsus appliance" OR "Twin force bite corrector" OR "Eureka spring" OR "Sabbagh spring" OR Activator OR Bionator OR "Bimler appliance" OR "Fraenkel appliance" OR "Frankel appliance" OR "Bass appliance" OR "Harvold appliance" OR "Andresen appliance" OR "Teuscher appliance" OR "Stoeckli appliance" OR "Stockli appliance" OR Biobloc OR "Bite jumper" OR "Bite jumping" OR "SII appliance" OR "Twin block") AND ("class ii malocclusion" OR class ii div* OR class ii/ OR ("class ii" AND orthodont*))	*No limitations*	**959**

Google Scholar Searched on October 10, 2014 http://scholar.google.gr/advanced_scholar_search?hl=en&as_sdt=0,5	functional AND orthopedic AND orthodontic AND treatment AND malocclusion AND appliance AND appliances AND “class ii”	*in Medicine, Pharmacology, and Veterinary Science*	**1047**
Web of Science Searched on October 10, 2014 http://apps.webofknowledge.com/WOS_GeneralSearch_input.do?last_prod=WOS&SID=P1G3aMpjDambMDFjp3e&product=WOS&highlighted_tab=WOS&search_mode=GeneralSearch	TS=(maxill* AND (excess* OR prognath*)) OR TS=(mandib* AND (deficien* OR retrognath* OR reposition* OR enhanc* OR advanc*)) OR TS=(functional) OR TS=(orthopaedic*) OR TS=(orthopedic*) OR TS=(growth) OR TS=("Mandibular anterior repositioning appliance") OR TS=("Mandibular protraction appliance") OR TS=("Ritto appliance") OR TS=("Biopedic appliance") OR TS=("Mandibular corrector appliance") OR TS=("Mandibular advancing repositioning splint") OR TS=("Ventral telescope") OR TS=("Magnetic telescopic device") OR TS=(Herbst) OR TS=("Fraenkel appliance") OR TS=("Bimler appliance") OR TS=(Bionator) OR TS=(Activator) OR TS=("Sabbagh spring") OR TS=("Eureka spring") OR TS=("Twin force bite corrector") OR TS=("Forsus appliance") OR TS=("Forsus device") OR TS=("Forsus nitinol flat spring") OR TS=("Bite fixer") OR TS=("Adjustable bite corrector") OR TS=("Flex developer") OR TS=("Scandee tubular jumper") OR TS=("Jasper jumper") OR TS=("Functional mandibular advancer") OR TS=("MARA") OR TS=("Twin block") OR TS=("SII appliance") OR TS=("Bite jumping") OR TS=("Bite jumper") OR TS=(Biobloc) OR TS=("Stockli appliance") OR TS=("Stoeckli appliance") OR TS=("Teuscher appliance") OR TS=("Andresen appliance") OR TS=("Harvold appliance") OR TS=("Bass appliance") OR TS=("Frankel appliance") AND (TS=("class ii malocclusion") OR TS=(class ii div*) OR TS=(class ii/) OR TS=("class ii" AND orthodont*))	in *“Dentistry Oral Surgery Medicine”*	**1191**
Evidence-Based Medicine Searched on October 10, 2014 http://ebm.bmj.com/searchall/	functional AND "class ii" AND orthodont*	functional AND "class ii": *in title or abstract* orthodont*: *anywhere in article*	**62**
Scopus Searched on October 10, 2014 http://www.scopus.com/search/form.url?display=advanced&clear=t&origin=searchbasic&txGid=1lkb0B3HcbSzUk8cVtIzKL_%3a3	(Herbst OR "Magnetic telescopic device" OR "Ventral telescope" OR "Mandibular advancing repositioning splint" OR "Mandibular corrector appliance" OR "Biopedic appliance" OR "Ritto appliance" OR "Mandibular protraction appliance" OR "Mandibular anterior repositioning appliance" OR "MARA" OR "Functional mandibular advancer" OR "Jasper jumper" OR "Scandee tubular jumper" OR "Flex developer" OR "Adjustable bite corrector" OR "Bite fixer" OR "Forsus nitinol flat spring" OR "Forsus device" OR "Forsus appliance" OR "Twin force bite corrector" OR "Eureka spring" OR "Sabbagh spring" OR Activator OR Bionator OR "Bimler appliance" OR "Fraenkel appliance" OR "Frankel appliance" OR "Bass appliance" OR "Harvold appliance" OR "Andresen appliance" OR "Teuscher appliance" OR "Stoeckli appliance" OR "Stockli appliance" OR Biobloc OR "Bite jumper" OR "Bite jumping" OR "SII appliance" OR "Twin block") AND ("class ii malocclusion" OR class ii div* OR class ii/ OR ("class ii" AND orthodont*)) AND (functional OR orthopaedic* OR orthopedic* OR growth) AND (LIMIT-TO(SUBJAREA, "DENT") OR LIMIT-TO(SUBJAREA, "MULT"))	*in Dentistry*	**1749**
LILACS database Searched on October 10, 2014 http://bases.bireme.br/cgi-bin/wxislind.exe/iah/online/?IsisScript=iah/iah.xis&base=LILACS&lang=i	orthodont$ or angle class ii and functional	*in Words*	**389**
Bibliografia Brasileira de Odontolgogia Searched on October 10, 2014 http://bases.bireme.br/cgi-bin/wxislind.exe/iah/online/?IsisScript=iah/iah.xis&base=BBO&lang=i	orthodont$ or angle class ii and functional	*in Words*	**219**
Ovid database Searched via HEAL-Link on October 10, 2014 http://ovidsp.tx.ovid.com/sp-3.4.2a/ovidweb.cgi?&S=LCNCFPPJBPDDFDPJNCBLAFFBMDFPAA00&New+Database=Single%7c2	((maxill* and (excess* or prognath*)) or (mandib* and (deficien* or retrognath* or reposition* or enhanc* or advanc*)) or functional or orthopaedic* or orthopedic* or growth or (Herbst or "Magnetic telescopic device" or "Ventral telescope" or "Mandibular advancing repositioning splint" or "Mandibular corrector appliance" or "Biopedic appliance" or "Ritto appliance" or "Mandibular protraction appliance" or "Mandibular anterior repositioning appliance" or "MARA" or "Functional mandibular advancer" or "Jasper jumper" or "Scandee tubular jumper" or "Flex developer" or "Adjustable bite corrector" or "Bite fixer" or "Forsus nitinol flat spring" or "Forsus device" or "Forsus appliance" or "Twin force bite corrector" or "Eureka spring" or "Sabbagh spring" or Activator or Bionator or "Bimler appliance" or "Fraenkel appliance" or "Fränkel appliance" or "Frankel appliance" or "Bass appliance" or "Harvold appliance" or "Andresen appliance" or "Teuscher appliance" or "Stoeckli appliance" or "Stockli appliance" or Biobloc or "Bite jumper" or "Bite jumping" or "SII appliance" or "Twin	*in title, abstract, full text, caption text*	**276**

	block")).mp. and (("class ii malocclusion" or class ii div*).mp. or class ii/ or ("class ii" and orthodont*).mp.)		
Bandolier Searched on October 10, 2014 http://www.medicine.ox.ac.uk/bandolier/bformHJ.html	orthodontic	No limitations	**1**
Atypon Link Searched on November 12, 2011***** http://www.atypon-link.com/	orthodont*	No limitations	**16**
African Journals Online Searched on October 10, 2014 http://www.ajol.info/index.php/index/search	(functional OR class ii) AND orthodont*	No limitations	**6**
ProQuest Searched on October 10, 2014 http://proquest.umi.com/pqdweb?RQT=403&TS=1321887206&clientId=68919	(functional) AND (orthodont*) AND ("class ii")	*in Anywhere except full text – ALL*	**123**
Conference Paper Index Searched via Cambridge Scientific Abstracts (1919 – October 10, 2014) http://journals.cambridge.org/action/advanceSearch?sessionId=30DBC2BF08FD26DF4AA4689F04961599.journals	(functional OR "class ii") AND orthodont*	*in Medicine*	**169**
German National Library of Medicine (ZB MED) Searched via MEDPILOT on October 10, 2014 http://www.medpilot.de/app/65de19b10e83c5faea497930a5cfb322?LANGUAGE=en	(orthodont* OR "class ii") AND appliance*	*in "Catalogue Medicine. Health."*	**198**
***meta*Register of Controlled Trials** Searched on October 10, 2014 http://www.controlled-trials.com/mrct/searchform	(functional OR "class ii") AND orthodont*****	*in "all registers"*	**32**
Sum			**9107**

*As of January 2012, Atypon Link is discontinued.

Supplementary table 2 A classification of the angular cephalometric variables examined in the present study, with the various terms used among the selected studies, and their definitions.

Term used	*Other terms*	*Definition*
	Angular skeletal cephalometric variables	
SNA	Sella-Nasion-Point A	The posterior-inferior angle formed by Se, Na and A point
SNB	Sella-Nasion-Point B	The inferior-posterior angle defined by Se, Na and B point
SNPg	-	The angle formed by Se, Na and Pg
ANB	A Point-Nasion-B Point angle, A-B Difference, A-B diff	The angle between Point A, Na and Point B; the difference between angles SNA and SNB
NAPg	NAP, Angle of convexity	The inferior angle formed by the lines Nasion-A point and A point-Pg
SGo:Nme (%)	Jarabak ratio, Sgo:NM, Sgo:Nme	Ratio of posterior to anterior facial height
SN-ML	SN-MP, SN-GoGn, SN-GoMe, ML/NSL, MPA	The angle formed by the anterior cranial base plane and the mandibular plane
NL-ML	Maxillary Mandibular Planes angle, ANSPNS-GoGn, MM angle, PP-MP, PP-GoMe, PP-GoGn, ANSPNS-GoMe, SpaSpp-Mgo	The angle formed by the Palatal line and the Mandibular line
SN-NL	Sella Nasion to Palatal Plane, SeNa-ANSPNS, NL/NSL, SN-ANS/PNS, SN-SpaSpp, SN-PP, PP angle, Palatal plane angle	The angle between the Sella-Nasion line and the Palatal plane
SN-OP	OL/NSL, OP angle, Occlusal plane angle	Angle between Sella-Nasion line and Occlusal plane
y axis		Angle between Frankfurt Horizontal Plane and S-Gn plane
	Angular dental cephalometric variables	
1s-SN	-	Angle from the upper incisor long axis to the anterior cranial base plane
1i-ML	1i-MP, IMPA, Incisor-mandibular plane angle, L1GoGn, L1ML, L1MP	Angle between lower incisor long axis and mandibular plane
1s-1i	Interincisal angle, L1U1, 1/1, 1iW, 1s/1i	The angle formed by the long axis of the upper and lower incisors
1s-NA	1NA, 1/NA, max. central incisor.NA	Angle between the upper incisor long axis and the NA line
1i-NB	L1NB, mand. central incisor.NB	Angle between lower incisor long axis and the NB line
1i-VL		Angle between the upper incisor long axis and a reference line (VL) vertical to the Frankfurt Horizontal plane.
	Angular soft tissue cephalometric variables	
N´SnPg´	Convexity angle excluding the nose	The angle formed by N', Sn and Pg' points
Nasolabial angle	CmSnLs	Angle between the nasal line and the most anterior point of the upper lip
Mentolabial angle	Li-Sm-Pg'	Angle between the line of the lower lip and the Me'
H angle	Holdaway Angle, Na'-Pg'-Ls, H line-Na'Pg'	The superior-posterior angle between the H line (the most anterior superior point of the soft tissue chin and Labrale superius) and the N'-Pg' line
	Ratios	
ANSMe:Nme	LFH/TAFH	Ratio of lower facial height to total anterior facial height
Gonial Ratio	-	The ratio between the upper gonial (NGoAr) and the lower gonial (NGoMe) angle
S-Ar/Ar-Go	-	Quotient of Ar-Go and S-Ar lines

Supplementary table 3a Assessment of randomized controlled trials with the Cochrane risk of bias tool.

A/A	Author/year	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of outcome assessment (detection bias)*	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias	Overall risk of bias
1	Baysal and Uysal (58)	Low risk	Unclear	Unclear	High risk	Low risk	Low risk	**High risk**

*The domain "blinding of participants and personnel" was not assessed, as it was judged unfeasible.

Supplementary table 3b Risk of bias for the non-randomized trials, according to the modified Downs and Black tool.

		1	2	3	4	5	6	7	8
	Author/year	Alali (57)	de Almeida et al, (59) and de Almeida et al, (60)	Gunay *et al.* (61)	Karacay *et al.* (62)	Latkauskienė (63)	Oztoprak *et al.* (64)	Phelan *et al.* (65)	Uyanlar *et al.* (66)
Reporting	Is the hypothesis/aim/objective of the study clearly described?	1	1	1	1	1	1	1	1
	Are the main outcomes to be measured clearly described in the Introduction or Methods section?	1	1	1	1	1	1	1	1
	Were inclusion and exclusion criteria clearly stated?	1	1	1	1	1	1	1	1
	Are the characteristics of the patients included clearly described?	1	1	1	1	1	1	1	1
	Are the functional appliances used clearly described?	1	1	1	1	1	1	1	1
	Are the distributors of principal confounders in each group of subjects to be compared clearly described?	1	1	0	0	0	0	0	0
	Are the main findings of the study clearly described?	1	1	1	1	1	1	1	1
	Does the study provide estimates of the random variability in the data for the main outcomes?	1	1	1	1	1	1	1	1
	Have all important adverse events that may be a consequence of functional appliances been reported?	1	1	1	1	1	1	1	1
	Have the characteristics of patients lost to follow-up been described?	1	1	0	0	1	0	1	0
	Have actual probability values been reported for the main outcomes except where the probability value is less than 0.001?	1	0	0	1	1	1	0	0
External validity	Were the patients asked to participate in the study representative of the entire population from which they were recruited?	1	1	1	1	1	1	1	1
	Were those subjects who were prepared to participate representative of the entire population from which they were recruited?	1	1	1	1	1	1	1	1
	Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?	1	1	1	1	1	1	1	1
Internal validity - bias	Was an attempt made to blind those measuring the main outcome of the intervention?	1	0	0	0	1	0	0	0
	If any of the results of the study were based on “data dredging”, was that made clear?	1	1	1	1	1	1	1	1
	Do the analyses adjust for different lenghts of follow-up of patients?	1	0	0	0	1	0	0	0
	Were the statistical tests used to assess the main outcomes appropriate?	1	1	1	1	1	1	1	1
	Was compliance with the functional appliance used reliable?	1	1	1	1	1	1	1	1
	Were the main outcome measures used accurate (valid and reliable)?	1	1	1	1	1	1	1	1
Internal validity - confounding	Were the patients in different intervention groups recruited from the same population?	1	0	0	1	0	1	0	0
	Were the baseline characteristics comparable?	1	1	0	0	1	1	1	1
	Were study subjects in different intervention groups recruited over the same period of time?	1	0	0	1	0	0	0	0
	Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	1	0	0	1	0	0	0	0
	Were losses of patients to follow-up taken into account?	0	0	0	0	0	0	0	0
Power	Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?	3	5	3	3	5	5	5	3
	Sum	**27**	**23**	**18**	**22**	**25**	**23**	**22**	**19**

Supplementary table 4 Details for the GRADE assessment of the primary outcomes*. SGs, subgroups; OIS, optimal information size.

Variable	Risk of bias	Inconsistency	Indirectness	Imprecision	Reporting biases	Effect magnitude	Dose response effect	Residual confounding
SNA	No clear indication to downgrade.	Low heterogeneity; no reason to downgrade	Directly relevant	Adequate sample	No evidence of bias	No reason to rate up	No significant regression to treatment duration	Existing
SNB	No clear indication to downgrade.	Moderate heterogeneity, explained by SGs and confidence regarding decision unaffected	Directly relevant	Adequate sample	No evidence of bias	No reason to rate up	No significant regression to treatment duration	Existing
ANB	No clear indication to downgrade.	High heterogeneity, explained by SGs; confidence regarding decision might be affected	Directly relevant	Adequate sample	No evidence of bias	No reason to rate up	No significant regression to treatment duration	Existing
SN-ML	No clear indication to downgrade.	Low heterogeneity; no reason to downgrade	Directly relevant	Adequate sample	No evidence of bias	No reason to rate up	No significant regression to treatment duration	Existing
1s-SN	No clear indication to downgrade.	High heterogeneity, explained by SGs and confidence regarding decision unaffected	Directly relevant	Adequate sample	No evidence of bias	Rated up by two for very large effect, which was included in the mean effect, the confidence interval and the prediction interval, while no serious limitations were found.	Significant regression of incisor inclination to treatment duration	Existing
1i-ML	No clear indication to downgrade.	High heterogeneity, explained by SGs; confidence regarding decision might be affected	Directly relevant	Adequate sample	No evidence of bias	Rated up by two for very large effect, which was included in the mean effect, the confidence interval and the prediction interval, while no serious limitations were found.	Significant regression of incisor inclination to treatment duration	Existing
Nasolabial angle	No clear indication to downgrade.	Low heterogeneity; no reason to downgrade	Directly relevant	Inadequate sample (210 patients) compared to the OIS (240 patients); confidence intervals includes the null-effect, but not large benefit or harm. No clear indication of imprecision.	No evidence of bias	No reason to rate up	No significant regression to treatment duration	Existing

*Regarding risk of bias, all outcomes start from "low", due to the inclusion of non-randomized studies; no reason to further downgrade.

Supplementary table 5 Sensitivity analysis by meta-regression on the duration of functional appliance treatment.

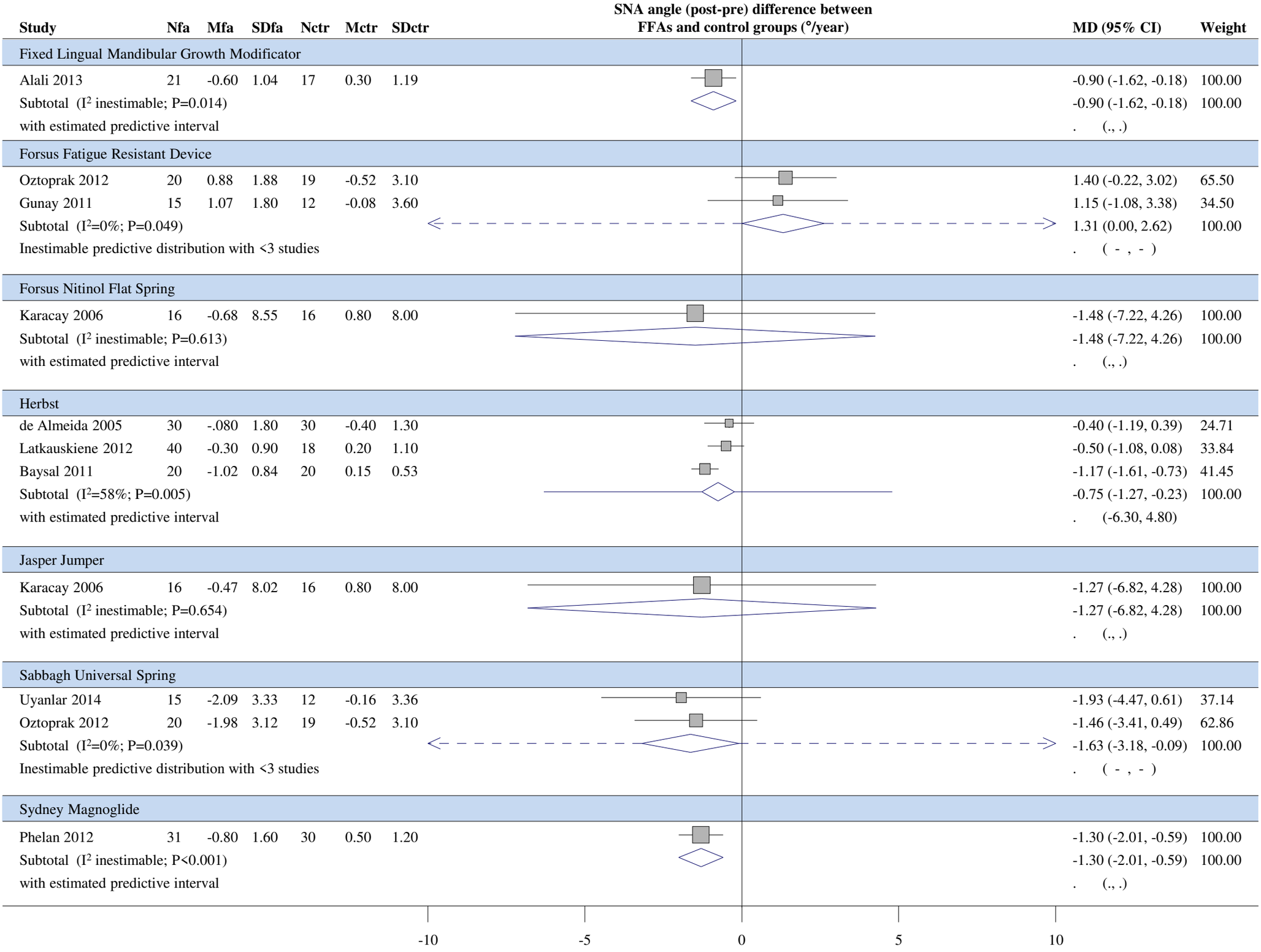
Variable	P value*
SNA	0.251
SNB	0.503
ANB	0.756
SN-ML	0.689
1s-SN	**0.054**
1i-ML	**0.059**
Nasolabial angle	0.832

*From meta-regression. Bold values indicate statistically significant results.

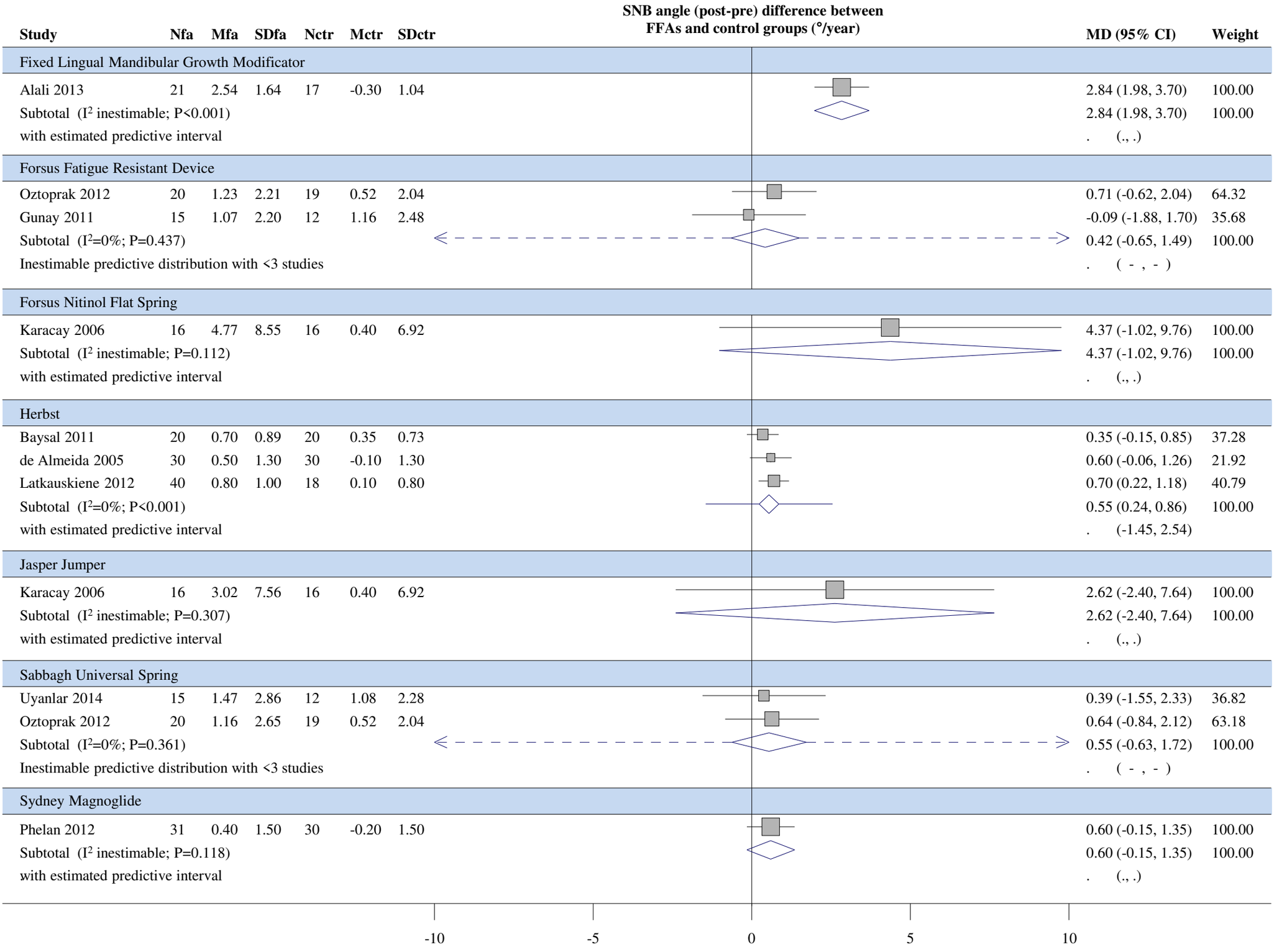
Supplementary table 6 Sensitivity analysis according to the study design of included studies. RCT, randomized controlled trial; MD, mean difference; CI, confidence interval; P_{SG}, P value for differences between subgroups of RCTs and non-RCTs.

		RCTs			Non-RCTs			
Variable	n	n	MD (95% CI)	P value	n	MD (95% CI)	P value	P_{SG}
SNA	9	1	-1.17 (-1.61,-0.74)	<0.001	8	-0.71 (-1.09,-0.32)	<0.001	0.240
SNB	9	1	0.35 (-0.15, 0.85)	0.174	8	0.97 (0.30, 1.64)	0.004	0.522
ANB	9	1	-1.63 (-2.20, -1.06)	<0.001	8	-1.76 (-2.67, -0.86)	<0.001	0.948
SN-ML	8	1	0.56 (-0.23, 1.35)	0.165	7	0.45 (-0.08, 0.98)	0.097	0.824
1s-SN	6	1	-4.24 (-6.19, -2.30)	<0.001	5	-8.83 (-14.06, -3.61)	0.001	0.409
1i-ML	9	1	2.74 (1.33, 4.15)	<0.001	8	8.81 (3.14, 14.48)	0.002	0.578
Nasolabial angle	5	1	-2.15 (-6.26, 1.96)	0.305	4	0.94 (-1.79, 3.67)	0.500	0.307

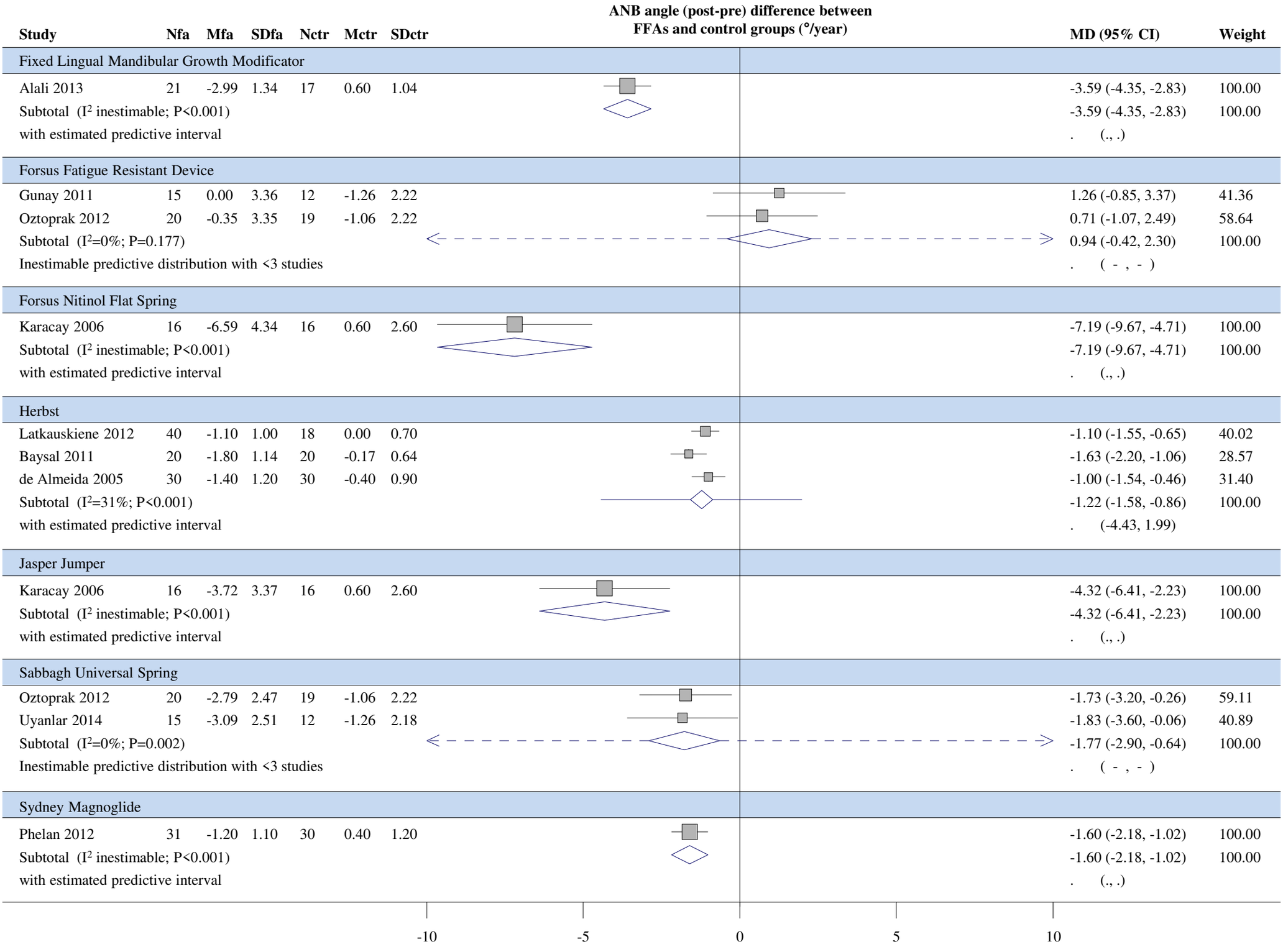
Supplementary figure 1 Meta-analysis of the SNA angle (in °/year) between the FFAs and control groups divided into subgroups according to the FFA used.



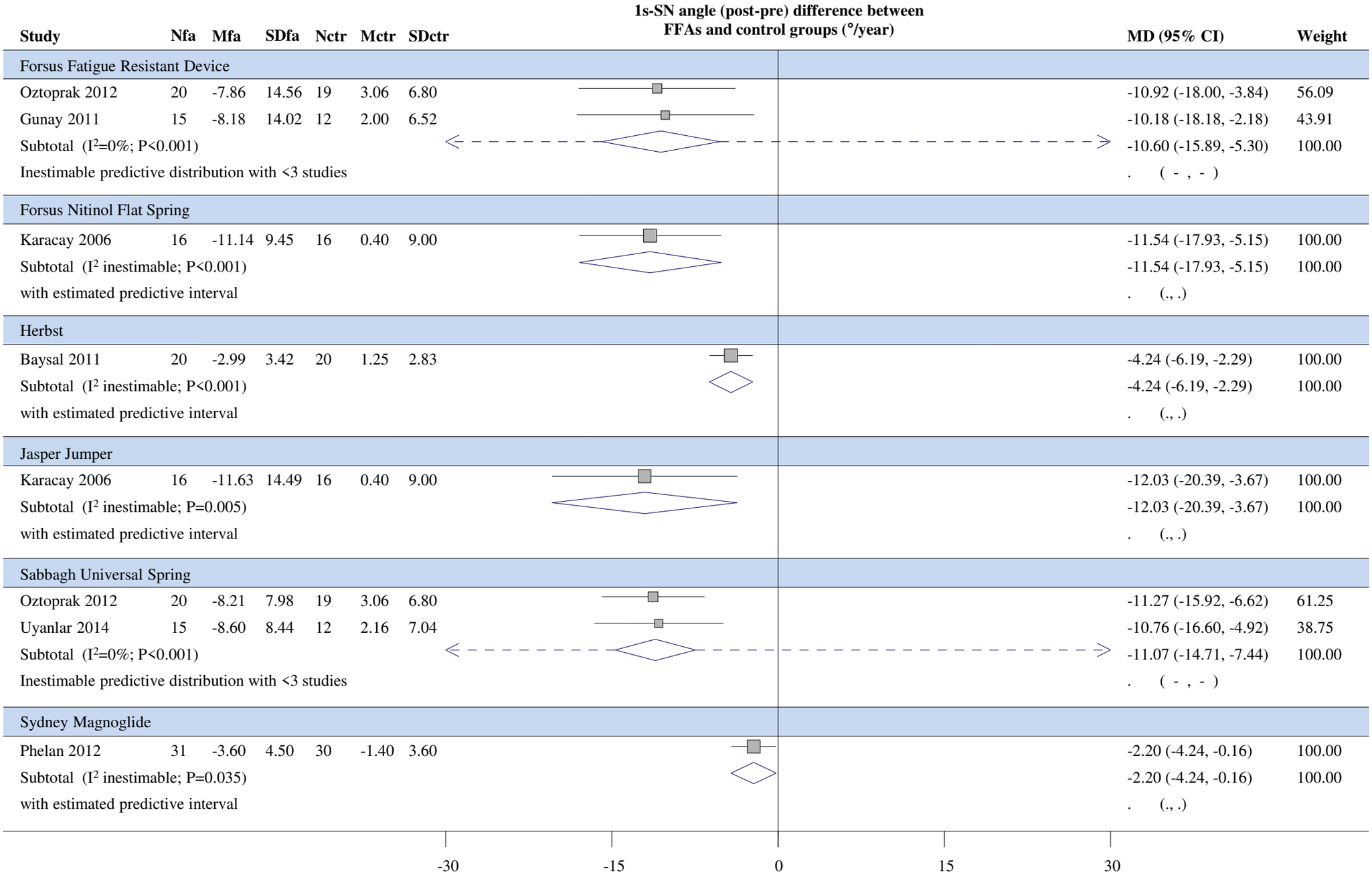
Supplementary figure 2 Meta-analysis of the SNB angle (in °/year) between the FFAs and control groups divided into subgroups according to the FFA used.



Supplementary figure 3 Meta-analysis of the ANB angle (in °/year) between the FFAs and control groups divided into subgroups according to the FFA used.



Supplementary figure 4 Meta-analysis of the 1s-SN angle (in °/year) between the FFAs and control groups divided into subgroups according to the FFA used.



Supplementary figure 5 Meta-analysis of the 1i-ML angle (in °/year) between the FFAs and control groups divided into subgroups according to the FFA used.